

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

DISTRIBUTIONAL IMPACT OF HEALTH CARE FINANCE IN SOUTH AFRICA

John Ele-Ojo Ataguba

Thesis presented for the degree of

DOCTOR OF PHILOSOPHY

in the

Department of Economics

Faculty of Commerce

University of Cape Town

November 2012

Supervisors

A/Prof. Ingrid Woolard (School of Economics)

Prof. Diane McIntyre (Health Economics Unit)

ABSTRACT

In South Africa, health care is financed through different mechanisms – allocations from general taxes, private health insurance contributions and direct out-of-pocket payments. These mechanisms impact differently on different households. While there is empirical evidence in developed countries, the distributional impact of such payments and methodological challenges in such assessments in the context of Africa are scarce. Borrowing from the tax literature, the thesis aims to assess the relative impact of health care financing on households' welfare and standards of living. Methodological issues around the assessment of income redistributive impact of health care payments in the context of South Africa are also explored.

Using the Income and Expenditure Survey 2005/06 data, and some plausible assumptions about who bears the final burden of health care payments, the results from the research indicate that total health care financing in South Africa is progressive (i.e. it absorbs a relatively smaller share of the poor's income compared to the corresponding shares among the richer groups). However, not all the constituent parts are progressive. Out-of-pocket payments are mildly regressive (i.e. poorer groups pay a higher proportion of their income than the richer groups). All indirect taxes are regressive and all direct taxes are progressive. Private health insurance was progressive but with a skewed distribution of enrolled members. These results are robust to the choice of equivalence scale parameters, shifting assumptions about corporate income tax, and the society's or policy maker's aversion to inequality. Further, financing health care through direct taxes, private health insurance and general taxes induce positive (or pro-poor) income redistributive effects (i.e. redistributing income from the rich to the poor) while indirect taxes and out-of-pocket payments induce negative (or pro-rich) income redistributive effect (i.e. redistributing income from the poor to the rich). The combined effect of these is a pro-poor income redistributive effect associated with total health care financing in South Africa. In all cases, the progressivity effect dominates the horizontal and reranking effects but this is less so for out-of-pocket payments and private health insurance where substantial differential treatments occur. These results are also robust to several assumptions about the values of some parameter within reasonable limits.

The overall assessment has shown that there is scope for promoting a more equitable health care financing system in South Africa. This can be achieved especially by increasing the government's commitment to the health sector and increased public funding.

TABLE OF CONTENTS

ABSTRACT	I
TABLE OF CONTENTS	II
LIST OF TABLES	VI
LIST OF FIGURES	VII
LIST OF ABBREVIATIONS	IX
DEDICATION	XI
ACKNOWLEDGEMENTS	XII
CHAPTER 1	1
GENERAL INTRODUCTION	1
1.1 AIM AND OBJECTIVES	3
1.1.1 Aim	3
1.1.2 Objectives	3
1.2 THE STRUCTURE OF THE REMAINING CHAPTERS	4
CHAPTER 2	6
SOUTH AFRICAN ECONOMY AND THE HEALTH SYSTEM	6
2.1 INTRODUCTION	6
2.2 OVERVIEW OF SOUTH AFRICA AND THE ECONOMY	6
2.3 HEALTH INDICATORS IN SOUTH AFRICA	12
2.4 BRIEF OVERVIEW OF THE SOUTH AFRICAN HEALTH SYSTEM	14
2.5 HEALTH CARE EXPENDITURE AND FINANCING IN SOUTH AFRICA	22
2.6 HEALTH CARE FINANCING CHALLENGES IN SOUTH AFRICA	26
2.7 CURRENT HEALTH CARE FINANCING REFORM IN SOUTH AFRICA	29
2.8 SUMMARY	31
CHAPTER 3	32
THEORETICAL FRAMEWORK AND LITERATURE REVIEW	32
3.1 INTRODUCTION	32
3.2 THEORETICAL REVIEW	32
3.3 ECONOMICS AND THE CONCEPTS OF EQUITY AND FAIRNESS	33
3.4 REVIEW OF THE BASIC THEORIES OF JUSTICE	38
3.4.1 Utilitarianism	38
3.4.2 Egalitarianism	41
3.4.3 Rawls' maximin principle of justice	43

3.4.4	Envy principle	45
3.4.5	Libertarian principle of justice	46
3.4.6	Nozick's Entitlement Theory	47
3.5	APPROACHES TO ASSESSING EQUITY IN HEALTH CARE FINANCING	51
3.6	REDISTRIBUTION AND HEALTH CARE FINANCING	56
CHAPTER 4		61
METHODOLOGICAL AND EMPIRICAL REVIEW OF LITERATURE		61
4.1	INTRODUCTION	61
4.2	MEASURING PROGRESSIVITY IN HEALTH CARE FINANCING	62
4.2.1	Computing the Gini and the concentration indices	68
4.3	EMPIRICAL REVIEW OF PROGRESSIVITY IN HEALTH CARE FINANCING	70
4.4	DECOMPOSING REDISTRIBUTIVE EFFECT OF HEALTH CARE PAYMENTS	77
4.4.1	Developments in assessing redistributive effects	79
4.5	EMPRICAL REVIEW OF REDISTRIBUTIVE EFFECT OF HEALTH CARE FINANCING	93
4.6	GAPS IN THE LITERATURE	97
CHAPTER 5		108
DATA AND METHODOLOGY OF DATA EXTRACTION		108
5.1	INTRODUCTION	108
5.2	DATA	108
5.3	MEASURING HOUSEHOLD INCOME	110
5.3.1	Alternative measures	110
5.3.2	Adjusting income for underreporting	111
5.3.3	Adjusting income using adult equivalence scales	113
5.4	INCIDENCE ASSUMPTIONS AND PROCEDURE FOR EXTRACTING HOUSEHOLD HEALTH CARE FINANCING CONTRIBUTIONS	116
5.5	ADJUSTMENTS AND BENCHMARKING OF EXTRACTED FINANCING COMPONENTS	129
CHAPTER 6		131
EMPIRICAL ASSESSMENT OF THE INCIDENCE OF HEALTH CARE FINANCING IN SOUTH AFRICA		131
6.1	INTRODUCTION	131
6.2	OVERVIEW OF THE METHODOLOGY	131
6.3	EMPIRICAL RESULTS OF THE INCIDENCE OF HEALTH CARE FINANCING IN SOUTH AFRICA	135

6.4	FORMAL ASSESSMENT OF THE PROGRESSIVITY OF HEALTH CARE FINANCING IN SOUTH AFRICA	144
6.4.1	Direct taxes	144
6.4.2	Indirect taxes	146
6.4.3	Overall health care financing	148
6.5	SENSITIVITY AND ROBUSTNESS ANALYSES OF THE INCIDENCE RESULTS	154
6.5.1	Robustness of the Kakwani index to changes in aversion to inequality parameter	154
6.5.2	Robustness of the Kakwani index to changes in equivalence scale parameters	159
6.6	DISCUSSION	164
6.7	CONCLUSION	170
CHAPTER 7		171
	DECOMPOSITION ANALYSIS OF THE INCOME REDISTRIBUTIVE EFFECT OF HEALTH CARE FINANCING IN SOUTH AFRICA	171
7.1	INTRODUCTION	171
7.2	OVERVIEW OF THE METHODOLOGY	171
7.3	THE METHODOLOGY OF THE AJL DECOMPOSITION	174
7.3.1	The overall redistribution	175
7.3.2	The AJL vertical component	175
7.3.3	The AJL re-ranking component	176
7.3.4	The AJL horizontal component	176
7.4	THE METHODOLOGY OF THE UL DECOMPOSITION	176
7.4.1	The UL vertical component	177
7.4.2	The UL horizontal component	177
7.4.3	The UL re-ranking component	177
7.5	THE METHODOLOGY OF THE DJA DECOMPOSITION	178
7.6	RESULTS OF EMPIRICAL ASSESSMENT OF THE REDISTRIBUTIVE EFFECT OF HEALTH CARE FINANCING IN SOUTH AFRICA	186
7.7	A COMPARATIVE ANALYSIS OF INCOME REDISTRIBUTIVE EFFECTS USING THE AJL, UL AND DJA MODELS	195
7.8	SENSITIVITY ANALYSES	208
7.9	DISCUSSION	215
7.10	METHODOLOGICAL DEBATES	220
7.11	CONCLUSION	223
CHAPTER 8		224
	IMPLICATIONS OF THE INCIDENCE AND INCOME REDISTRIBUTIVE EFFECT FOR EQUITY IN HEALTH CARE FINANCING IN SOUTH AFRICA	224
8.1	INTRODUCTION	224

8.2	IMPLICATIONS FOR EQUITABLE HEALTH CARE FINANCING IN SOUTH AFRICA	224
8.3	CONTRIBUTIONS OF THIS STUDY TO LITERATURE	230
8.4	LIMITATIONS OF THE STUDY	231
8.5	RECOMMENDATIONS FOR FUTURE RESEARCH	233
	REFERENCES	235

University of Cape Town

LIST OF TABLES

Table 2.1: Selected health indicators, South Africa.....	12
Table 2.2: Health services roles of different spheres of government.....	16
Table 4.1: Summary of Kakwani indices of progressivity across studies	98
Table 4.2: Redistributive effect associated with out-of-pocket payment	100
Table 4.3: Redistributive effect associated with direct taxes	101
Table 4.4: Redistributive effect associated with indirect taxes	102
Table 4.5: Redistributive effect associated with general taxes	103
Table 4.6: Redistributive effect associated with private health insurance	104
Table 4.7: Redistributive effect associated with social health insurance	105
Table 4.8: Redistributive effect associated with total public finance	106
Table 4.9: Redistributive effect associated with total private finance	106
Table 4.10: Redistributive effect associated with total health care financing	107
Table 5.1: Brief summary of the Income and Expenditure Survey, 2005/2006	109
Table 5.2: Income tax rates, rebates, and thresholds in South Africa (2005/06 assessment year)	119
Table 5.3: Herfindahl Index, Ratio for South Africa (2003 - 2006)	122
Table 5.4: Average cost of transportation and fuel consumption	126
Table 5.5: Summary of the computation methodology for each component of health care funding	127
Table 6.1: Cumulative shares of payments (direct taxes) by income quintile, South Africa 2005/2006	139
Table 6.2: Cumulative shares of payments (indirect taxes) by income quintile, South Africa 2005/2006	141
Table 6.3: Cumulative shares of health care payments by income quintile, South Africa 2005/2006	143
Table 6.4: Health care financing by source and Kakwani for total payments, South Africa 2005/2006	151
Table 6.5: Results of dominance tests, South Africa 2005/2006.....	152
Table 6.6: Cumulative shares of health care payments by income quintile, South Africa 2005/2006	153
Table 7.1: <i>DJA</i> decomposition of the income redistributive effect of health care payments, South Africa 2005/2006.....	188
Table 7.2: <i>DJA</i> decomposition of the income redistributive effect of health care payments, South Africa 2005/2006.....	194
Table 7.3: Income redistributive effect of direct taxes for health care financing, South Africa 2005/2006	197
Table 7.4: Income redistributive effect of indirect taxes for health care financing, South Africa 2005/2006	201
Table 7.5: Income redistributive effect of indirect taxes for health care financing, South Africa 2005/2006	203
Table 7.6: Income redistributive effect of general taxes, private health insurance and out-of-pocket payments for health care financing, South Africa 2005/2006	205
Table 7.7: Income redistributive effect of overall health care financing, South Africa 2005/2006.....	207
Table 7.8: <i>UL</i> decomposition of the re-ranking effects into between group, within group, and entire group re- ranking	208

LIST OF FIGURES

Figure 2.1: Human Development Index, South Africa, 2011	7
Figure 2.2: Real Gross Domestic Product (GDP) growth in South Africa, 1994-2010.....	8
Figure 2.3: Government debt as a proportion of Gross Domestic product (GDP), 1994-2009.....	10
Figure 2.4: Public sector expenditure shares 2000/01 and 2010/11	11
Figure 2.5: The distribution of ill-health and disability in South Africa, 2002 - 2008.	13
Figure 2.6: Average life expectancy, South Africa, 2001-2010	14
Figure 2.7: Health care spending as a proportion of a country's Gross Domestic Product (GDP)	18
Figure 2.8: Health care spending as a proportion of GDP for selected countries, 2009	19
Figure 2.9: Health care expenditure (public vs. private) in Africa, 2006	20
Figure 2.10: Health care expenditure (public vs. private) in selected countries, 2009	21
Figure 2.11: Flow of health care funds in the South African health system, 2008	24
Figure 2.12: Trends in real per capita spending by medical schemes and the public health sector	25
Figure 2.13: Total expenditure on health per capita in Africa, 2006 (in US\$)	27
Figure 2.14: Health care spending per capita for selected upper middle income countries, 2009	28
Figure 4.1: Measures of progressivity illustrated	67
Figure 4.2: A descriptive illustration of the components of redistributive effects.....	84
Figure 6.1: Progressivity and regressivity illustrated using the progressivity curve	134
Figure 6.2: Health care payments as a proportion of income, South Africa 2005/2006.....	136
Figure 6.3: Concentration curves of health care payments, South Africa	138
Figure 6.4: Progressivity curves for direct taxes.....	145
Figure 6.5: Progressivity curves for indirect taxes	147
Figure 6.6: Progressivity curves for general taxes, private health insurance, out-of-pocket payments and overall health care financing	149
Figure 6.7: Single parameter ethical weighting function ($\kappa(p; \nu)$)	156
Figure 6.8: The impact on progressivity of changes in the aversion to inequality parameter.....	158
Figure 6.9: The impact of the choice of equivalent scale parameters on the assessment of progressivity.....	161
Figure 6.10: The impact of the choice of equivalence scale parameters in ordering progressivity ($\nu = 2$).....	163
Figure 6.11: The sensitivity of the Kakwani index of total health care financing to assumptions about corporate income tax	164
Figure 7.1: Variations in groups of close-equals.....	175
Figure 7.2: Policy maker's degree of equality mindedness	181
Figure 7.3: The single parameter ethical weighting function ($w(p; \nu)$)	182
Figure 7.4: Density plot of income distribution in South Africa, 2006.....	196
Figure 7.5: <i>AJL</i> and <i>UL</i> decomposition of the redistributive effect of direct and indirect taxes	198
Figure 7.6: <i>AJL</i> and <i>UL</i> decomposition of the redistributive effect of health care financing	204

Figure 7.7: Sensitivity of the redistributive effect to choice of inequality aversion parameters under the decomposable Gini index	209
Figure 7.8: Sensitivity of the vertical effect of health care financing to choice of parameters using the <i>DJA</i> model	211
Figure 7.9: Sensitivity of the horizontal effect of health care financing to choice of parameters using the <i>DJA</i> model	212
Figure 7.10: Sensitivity of the re-ranking effect of health care financing to choice of parameters using the <i>DJA</i> model	213
Figure 7.11: Sensitivity of the redistributive effect of health care financing to choice of parameters using the <i>DJA</i> model	214

University of Cape Town

LIST OF ABBREVIATIONS

AE	Adult equivalence
AIDS	Acquired Immune Deficiency Syndrome
AJL	Aronson-Johnson-Lambert methodology
ANC	African National Congress
ART	Acute Respiratory Tract Infection
ATP	Ability to pay
BG	Between group
BIA	Benefit Incidence Analysis
BP	Blood pressure
COICOP	Classification of Individual Consumption According to Purpose
CPI	Consumer Price Index
DAD	A freely available software for distributive analysis
DALYs	Disability-adjusted life years
DASP	Distributive Analysis Stata Package
DJA	Duclos-Jalbert-Araar methodology
EG	Entire group
GDP	Gross Domestic Product
GEAR	Growth, Employment and Redistribution
GHI	Government Health Insurance
H	Horizontal effect
HDI	Human Development Index
HE	Horizontal equity
HHI	Hirschman-Herfindahl Index
HIV	Human Immunodeficiency Virus
IES	South African Income and Expenditure Survey
MCA	Multiple Comparison Approach
MDGs	United Nations Millennium Development Goals
NHI	National Health Insurance
NHIF	National Health Insurance Fund
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary least squares

OOP	Out-of-pocket
PAYE	Pay-as-you-earn
PSU	Primary sampling units
QALYs	Quality-adjusted life years
R	Reranking effect
RE	Redistributive effect
RS	Reynolds-Smolensky methodology
SITE	Standard Income Tax on Employees
STC	Secondary Tax on Companies
STD	Sexually Transmitted Disease
TB	Tuberculosis
UK	United Kingdom
UL	Urban-Lambert methodology
UNDP	United Nations Development Programme
US	United States
US\$	United States dollars
USA	United States of America
V	Vertical effect
VAT	Value-added tax
VCL	van de Ven-Creedy-Lambert methodology
VE	Vertical equity
WG	Within group
WHO	World Health Organisation
ZAR	South African rand

DEDICATION

To

Arebi and Ochege

University of Cape Town

ACKNOWLEDGEMENTS

I express my profound appreciation to my supervisors – Associate Professor Ingrid Woolard and Professor Di McIntyre who have been very helpful in providing overall guidance and support throughout this my research journey. You made the road less difficult for me. I also want to express my profound appreciation to Dr Arnab Acharya of the London School of Hygiene and Tropical Medicine for the supervisory support I received while I was doing my fellowship in London. The discussions and engagements we had were very stimulating. Your idea of putting everything on the white board is exciting. I also thank my mentor and friend Professor Gavin Mooney who took the pains to go through initial drafts of my thesis. You are a fine academic and I greatly acknowledge your mentorship.

I also thank the Commonwealth Scholarships for the financial support that enabled me to spend some time at the University of London while working on my thesis. Without this support, it would have been very difficult to complete my research in due time. I also acknowledge the European Commission grant (Sixth Framework Programme; Specific Targeted Research Project no: 32289) and the IDRC (Grant number 103457) to the SHIELD project from where this thesis was born.

My appreciation also goes to my colleagues at the Health Economics Unit (HEU), University of Cape Town under the leadership of Associate Professor Susan Cleary for the fruitful academic discussions and interactions. I really found a home at the HEU. I also want to specifically acknowledge my colleagues who were willing to take on extra duties during the period I was away.

I also acknowledge the support and assistance of my friends and members of the Strategies for Health Insurance for Equity in Less Developed Countries (SHIELD) research team. I particularly thank Drs James Akazili and Gemini Mtei who have been helpful at one point or another during my research journey. My appreciation also goes to my very good friend and spiritual director, Fr. Hyacinth Ichoku for your love and support during my studies.

I remain eternally grateful to my wife Ochea who had to bear some hardships during my period of studentship, particularly while I was away on a fellowship in London. I also appreciate your support for me when I spend my weekends working on my research. You were also supportive of my odd working hours during the week. I really appreciate you. I want

to also appreciate my lovely daughter Arebi for bearing my absence from the home during the weekends; and my parents and siblings for their love, prayers and support throughout the period of my research.

Last but not least, I acknowledge God Almighty for His immense love throughout the period of my studies. When this journey started, every step I take makes it seemed farther and I kept wondered when it will come to an end. God you are faithful.

The usual disclaimer applies and all errors remain my own.

University of Cape Town

Chapter 1

“He sat down opposite the treasury and observed how the crowd put money into the treasury. Many rich people put in large sums. A poor widow also came and put in two small coins worth a few cents. Calling his disciples to himself, he said to them, “Amen, I say to you, this poor widow put in more than all the other contributors to the treasury. For they have all contributed from their surplus wealth, but she, from her poverty, has contributed all she had, her whole livelihood.” **Mark 12:41-44**

GENERAL INTRODUCTION

The link between income levels and health has been recognised by scholars (Narayan *et al.*, 2000, Deaton, 2002, Wagstaff, 2002a, Deaton, 2003). It is the case that low income levels (poverty) and ill-health reinforce each other because “poverty breeds ill-health, and ill-health keeps poor people poor” (Wagstaff, 2002a p.97). Poor health outcomes affect income generating capacities of individuals. Also social protection in the forms of social insurance and safety nets impacts on the health of the population (MDG Report, 2011). The importance of a healthy population for the growth of an economy is therefore crucial (Grossman, 1972, van Zon and Muysken, 2005). On this issue van Zon and Muysken (2005 p.41) write: “[m]ore and more economists have come to recognise that the relationship between health and economic growth is not only demand driven, but that health itself is an important determinant of economic growth”. Also the World Health Organisation (WHO) notes that “[p]romoting and protecting health is essential to human welfare and sustained economic and social development” (World Health Organization, 2010 p. ix). Though improvements in population health cannot be attributed to the health sector alone (Commission on Social Determinants of Health, 2008), it is still the case that health sector reforms have significant impacts on the economy because it determines *inter alia* the health of the population. The role of the state in such reforms is also critical especially as it relates to securing social protection and trying to meet the expectations of the population (World Health Organization, 2010).

If health is a major determinant of economic growth and social development, it is important to assess the extent to which the health system is performing in ensuring a healthy population.

One of the key dimensions for assessing the performance of the health system, according to the WHO is the extent of 'fairness' of its financing system (World Health Organization, 2000). Policy makers within and across countries have been urged to devise ways to ensure such fairness in paying for health care (Wagstaff, 2002b) by building on existing health care financing structures. This is so because "[n]o country starts from scratch in the way it finances health services. All have some form of system in place and must build on it according to their values, constraints and opportunities" (World Health Organization, 2010 p.87).

While the challenge of ensuring fairness in financing health care is common to all countries – developing and developed, the series of studies that have investigated this issue extensively have been concentrated among developed countries (cf. Wagstaff *et al.*, 1989, Wagstaff and van Doorslaer, 1992, 1997, van Doorslaer *et al.*, 1999, Wagstaff *et al.*, 1999b). Recently, there has been a call for countries to ensure universal health coverage for their population (World Health Organization, 2010). This renewed interest has charged countries, particularly the developing countries, to provide evidence relating to the extent to which their health systems ensure fairness in financing health care. This is also the case with South Africa as the country is currently considering embarking on a major restructuring of its health system in the form of a National Health Insurance (NHI) (Department of Health, 2011) with the goal of providing universal financial protection. In this regard there is a need for evidence on the extent to which the current health care financing system ensures fairness in financing.

Though the issue of fairness is normative, economists have developed some analytical tools to assess it in the context of tax systems. These tools have been used to assess the incidence and distributional impacts of taxes (Kotlikoff and Summers, 1987, Fullerton and Metcalf, 2002). It is the case that taxes could alter prices and returns on factors of production (Kotlikoff and Summers, 1987) hence it leads to a change in the economy's equilibrium. As the field of public sector economics and finance investigating incidence has grown over the years, some economists have become increasingly interested in a specialised aspect of public economics that studies health care financing (see for example van Doorslaer *et al.*, 1999, Wagstaff *et al.*, 1999b). They have used similar tools developed and used in public economics to understand who pays for and who bears the burden of such health care payments. Just as "the main purpose of taxation is to raise resources to finance government expenditure" (Burgess and Stern, 1993 p. 762), health care financing is concerned mainly with how to raise resources to finance health care needs (Bilger, 2008).

Health care can be financed in various ways and the composition of these mechanisms varies from developed to developing countries. Some countries rely heavily on tax funded systems of health care financing (Xu *et al.*, 2003) while some are predominantly based on direct out-of-pocket payments (Ataguba, 2012b). In some other countries, taxes are combined with other mechanisms of financing such as out-of-pocket payments, private insurance, and other forms of prepayment for health care. In South Africa, health care is financed through allocations from the general tax pool (approx. 40%), medical scheme contributions (approx. 45%), and out-of-pocket payments (approx. 14%) (McIntyre *et al.*, 2007). To date, there has been no comprehensive assessment of the progressivity and the redistributive impact of these different financing mechanisms or the health care financing system overall. This thesis however attempts to fill this gap by providing empirical evidence, combined with exploring some key methodological debates, around such assessment in South Africa.

1.1 AIM AND OBJECTIVES

1.1.1 Aim

The thesis aims to assess, for the first time in South Africa, the distributional impact of health care financing and whether such financing is equitable. This is done by paying particular attention to the impact of financing on households' income.

1.1.2 Objectives

The major objectives of the thesis include:

1. Assessing the progressivity of each health care financing mechanism in South Africa;
2. Assessing the overall progressivity in the South African health care financing system;
3. Testing the sensitivity of the measure of progressivity of health care financing to the choice of policy maker's (or society's) aversion to inequality;
4. Assessing the extent of the income redistributive effect of each health care financing mechanism in South Africa;
5. Quantifying the extent of the income redistributive effect of South Africa's overall health care financing system;

6. Comparing the impact of the choice of methodologies for the assessment of the income redistributive effect of health care financing in South Africa;
7. Testing the sensitivity of the measure of income redistribution in health care financing in South Africa to the choice of policy maker's (or society's) aversion to inequality;
8. Assessing the implications of the results for equity in health care financing in South Africa.

1.2 THE STRUCTURE OF THE REMAINING CHAPTERS

In order to achieve these objectives, the thesis is divided into nine chapters. In Chapter 2 an overview of the South African economy and health system is presented. It also discusses key health care financing challenges. This chapter only contains a summary of issues that are particularly relevant for understanding the analyses contained in this thesis.

Chapter 3 provides a review the theoretical literature on equity in health care financing. This chapter is important because it reviews the major theories of justices which underpin equity evaluations. Given that the issues considered in this thesis are normative, these theories were discussed in the context of health care financing. The various approaches to assessing equity in health care financing were also reviewed. The chapter also reviews the literature on redistribution as it relates to health care financing. Chapter 4 continues the review of literature but focusses on the methodological debates in the assessment of the (re)distributional impact of health care financing. This is also accompanied by a detailed review of the empirical literature from around the world on the assessment of equity in health care financing.

Chapter 5 introduces and discusses the dataset used in this thesis. This includes the assumptions made in extracting the various components of health care financing. It explains how attempts were made to ensure that these assumptions are based on relevant theory and evidence.

Chapter 6 presents and discusses the empirical results of the analysis of the incidence of health care financing in South Africa. This begins by looking at the individual financing mechanisms through to the overall financing system. The methodologies used for such assessments are also contained in this chapter. Sensitivity and robustness analyses were performed in order to assess the impact of changing the assumptions about the economies of scale parameters and the

choice of the policy maker's aversion to inequality on the measures of progressivity. The results obtained here were then compared with those from similar studies around the world.

Chapter 7 contains the empirical assessment of the income redistributive effect of health care financing in South Africa. Individual health care financing mechanisms were assessed as well as the overall health care financing system. This chapter unpacks the various components of income redistribution so as to identify the areas of health care financing in South Africa that drive inequity in the system. Sensitivity and robustness analyses were also performed to assess the impact of changing the assumptions about certain parameters on the overall results. From a methodological perspective, this chapter compares the results from competing models and draws some methodological conclusions in the case of South Africa, a country with a high income inequality.

In Chapter 8 the results from preceding chapters are discussed, paying particular attention to the key issues arising from such analyses. It attempts to provide recommendations on how to ensure a progressive and pro-poor South African health care financing system. In conclusion, this chapter also provides some recommendations for future work, limitations of the study and summarises the major contributions of the thesis to the body of knowledge.

Chapter 2

SOUTH AFRICAN ECONOMY AND THE HEALTH SYSTEM

2.1 INTRODUCTION

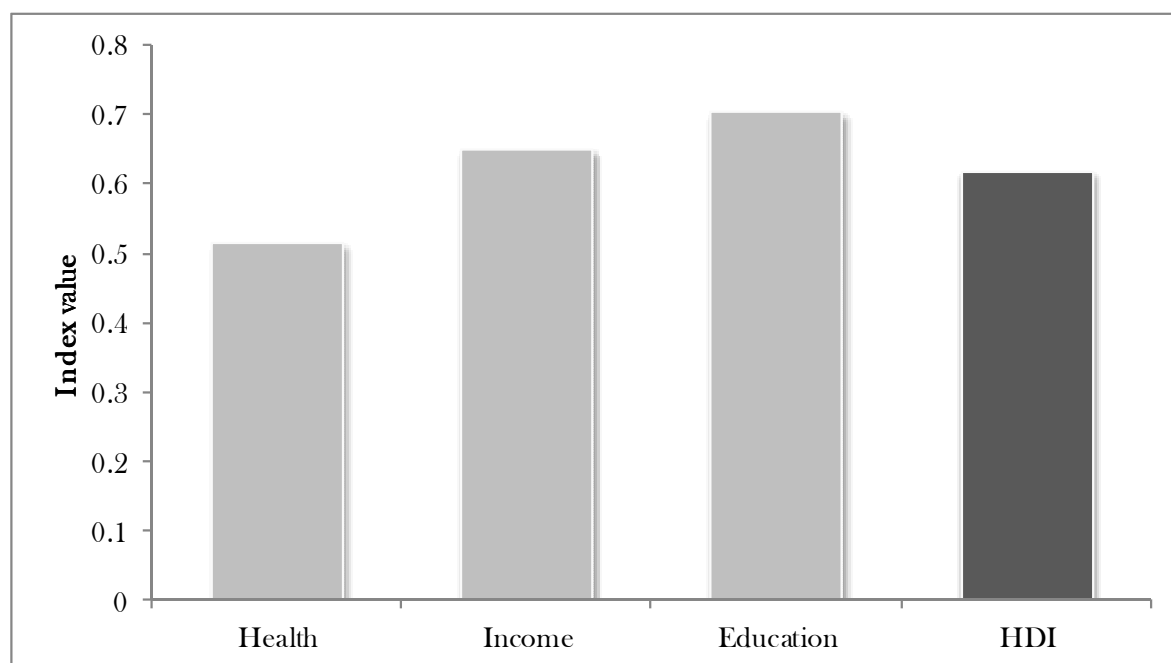
This chapter provides a brief overview of the South African economy, an overview of the population's health and an overview of the overall health system. It provides an explanation of the health care financing arrangements in the country and points to some of the challenges faced by the health care financing system. This overview forms a background that will enable the reader to gain a snapshot understanding of the economy and the health system upon which this thesis is based.

2.2 OVERVIEW OF SOUTH AFRICA AND THE ECONOMY

South Africa has an estimated population in 2011 of about 50.6 million people (Statistics South Africa, 2011) of whom more than 60% live in urban areas. Between 2008 and 2011, the population growth rate was estimated to be 1.1% (Statistics South Africa, 2011, World Bank, 2011a) making it one of the lowest in Africa. The male population grew slightly faster than the female population possibly due to the higher burden of HIV/AIDS among the female population (Cleary *et al.*, 2011). It had also been previously reported that the ratio of the number of male births to that of female births in South Africa is significantly greater than 1 (Garenne, 2004). Other African countries with population growth rates lower than that of South Africa include Mauritius (0.5%), Zimbabwe (0.5%) and Lesotho (0.8%) (World Bank, 2011a). South Africa has a Human Development Index (HDI) of 0.619 and is ranked 123rd out of 187 countries (UNDP, 2011). As shown in Figure 2.1, of the three constituents of the HDI, the contribution of health to the index is the lowest at 0.52 while education is the largest contributor. The HDI has continued to decline from a higher value of 0.724 in 1995 (shortly after democracy) to 0.695 in 2000 (du Toit, 2002) and to its current value in 2011. This

reflects a decline in the overall level of human development in the country post-independence. This decline is generally attributed to decreasing life expectancy associated with the increasing burden of HIV pandemic (du Toit, 2002, Cleary *et al.*, 2011, Kahn, 2011).

Figure 2.1: Human Development Index, South Africa, 2011

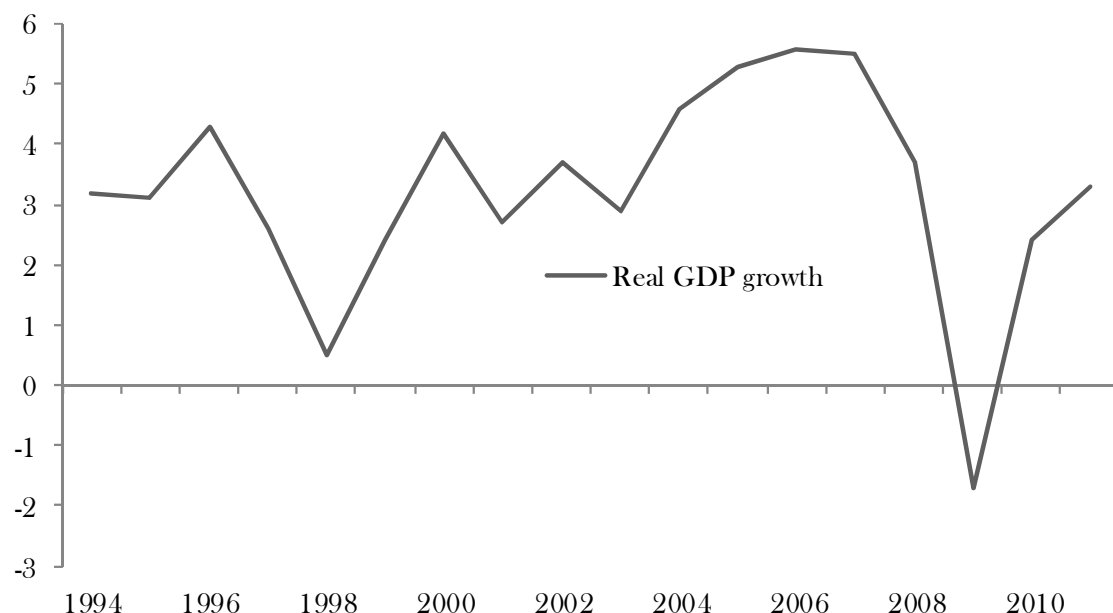


Source: UNDP (2011)

South Africa is a member of the Southern African Development Community (SADC) currently comprising 15 African countries. This is a regional integration and development organisation that focuses on the welfare of member states. South Africa is one of the few African countries to be categorised as an upper middle income country. Real gross domestic product (GDP) as at 2010 was US\$ 187 billion (in 2000 prices). While the nominal GDP per capita in 2010 was estimated at US\$ 7,275, per capita GDP at 2005 prices was estimated at US\$ 9,476 (in purchasing power parity terms). In 2010 the service sector, manufacturing sector and agriculture sector accounted for about 50%, 15% and 3% of the country's total GDP respectively (National Treasury, 2011). Similar patterns have existed in the past (National Treasury, 2004, 2005, 2006, 2007, 2008, 2009a, 2010). Between 1999 and 2008 the economy witnessed positive real per capita GDP growth (Development Indicators, 2010). Similarly real GDP (see Figure 2.2) showed positive growth from the period of transition to democracy (1994). The negative growth has been attributed to decline in mining and agricultural activities which were due largely to recent global financial crisis (African Economic Outlook, 2010b). The 2008 energy crisis in addition to a drop in the demand for exports contributed to the decline in mining activities in the country. Growth in the South African economy rebounded

in 2010 with real GDP growing at 2.4%. The fluctuating economic fortune has implications for the extent to which the government can spend on social services including health care.

Figure 2.2: Real Gross Domestic Product (GDP) growth in South Africa, 1994-2010



Source: Development Indicators (2010) for 1994-2009; African Economic Outlook (2010b) for 2010

For several decades, unemployment and poverty figures have been generally high (Klasen and Woolard, 2009). The latest data show that broad unemployment stood at 35.9% as at June 2010 with a greater proportion of unemployed women compared to unemployed men (Republic of South Africa, 2010). The high unemployment rates result from deteriorating labour market conditions with increasing decline in formal sector employment. For example, the labour force absorption rate declined from 56.3% in the 1990s to about 40.8% in 2010 (du Toit, 2002, Department of Labour, 2010). The deteriorating conditions of the labour market have been attributed *inter alia* to the recent economic downturn and a decline in the ratio of capital formation to GDP. These led to reductions in ability of the country to create new jobs, slow real economic growth in the face of increasing population, increasing demand for high skilled labour in the presence of oversupply of lower skilled labour, net emigration of skilled labour, strong presence of labour unions, and other labour legislations that hinder effective job creation (du Toit, 2002, Department of Labour, 2010).

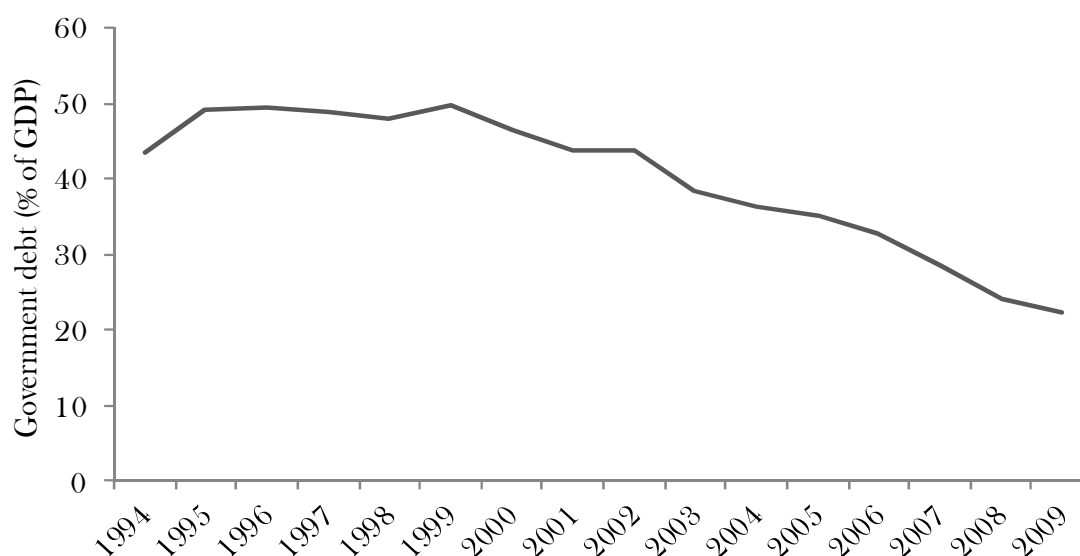
Using a national poverty line of R283 per person / month in 2008 prices, the proportion of poor people increased from 31% (in 1995) to 38% (in 2000). By 2005 it was estimated that about 23% of South Africans live below that poverty line (Development Indicators, 2010). A

greater proportion of blacks are categorised poor than the non-black population. In 2000 for instance, about 60% of blacks are categorised as poor compared to 20%, <10% and <5% of coloureds, Asians, and whites respectively (du Toit, 2002).

The period of apartheid was characterised by racial segregation and economic policies that favoured a selected minority of the population. This greatly contributed to the existence of economic and social inequalities. Post-apartheid (after 1994), South Africa was committed to addressing and redressing existing challenges that were carried over to the democratic government. Macroeconomic policy was backed with the commitment to ensure open markets, privatisation and creating an enabling environment for businesses and foreign investment (Weeks, 1999, McIntyre *et al.*, 2007). All these arose because of the need of the post-apartheid government to: increase the rate of economic growth and investment so as to modernise production; and to aim to achieve these increased rates in a way that would create wage employment, bring better wages and ensure greater equality in income and wealth distribution (Weeks, 1999). In 1996 the government adopted a strategy that ensured a reduction in the budget deficit (which avoids imposing a greater tax burden on households), tight monetary policy and trade liberalisation (Weeks, 1999, McIntyre *et al.*, 2007). The strategy was termed Growth, Employment and Redistribution (GEAR) and was based on neoliberal ideology that emphasises the market system. The GEAR policy (1996-2000) aimed to create employment and lead to substantial poverty reduction. However it did not succeed in delivering on its major objectives. The lack of success was blamed among other things, on the large fiscal contraction and high interest rates that discourage investments (Weeks, 1999).

One of the few successes recorded was in debt management. Government debt as a proportion of GDP, as shown in Figure 2.3 fell from 49% (in 1995) to 22% in 2009 (see McIntyre *et al.*, 2007). The decline in debt-to-GDP ratio as well as a reduction in government borrowing resulted mainly from the privatisation of state assets and cuts in state spending, improved revenue collection (including tax), and higher levels of economic growth (du Toit, 2002).

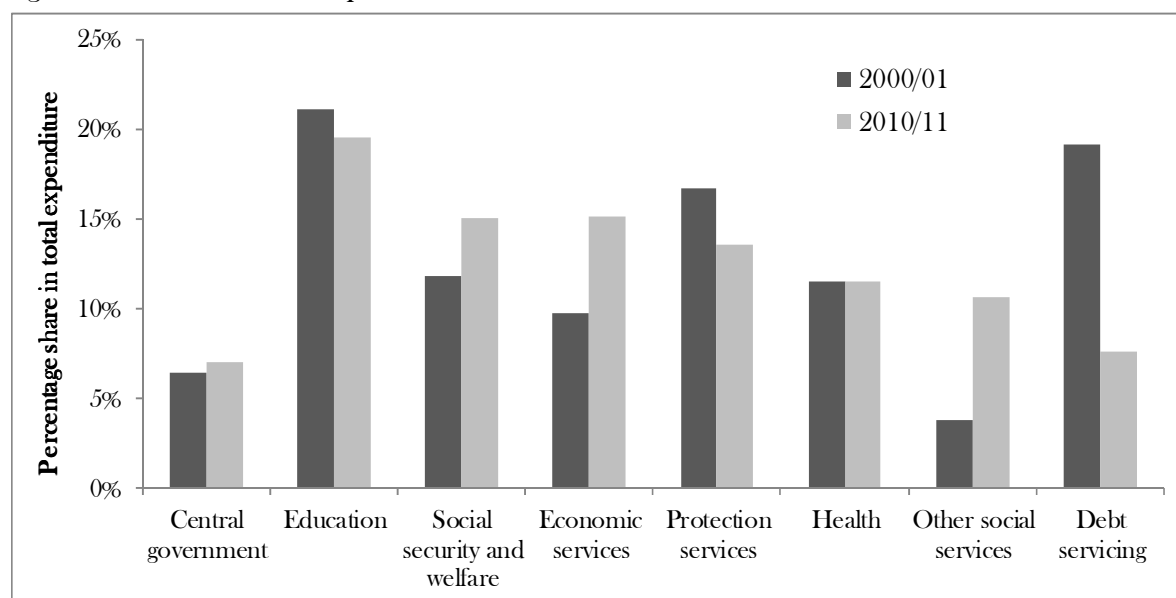
Figure 2.3: Government debt as a proportion of Gross Domestic product (GDP), 1994-2009



Source: Development Indicators (2010)

As the country continued to witness changing economic fortunes and government priorities, the shares of government expenditure on different activities also changed. As shown in Figure 2.4 some categories have witnessed increases in expenditure while others have experienced declines. While sectors such as social security and welfare, and economic services (transport and logistics, energy, water and sanitation services, etc.) experienced increases in expenditure share, education and protection services experienced decreases. The share of government expenditure on health remained relatively constant between 2000/01 and 2010/11. However it decreased between 2000/01 and 2007/08 (McIntyre *et al.*, 2007). The reduction in the share of education (and the constant share of health) is mainly due to the increased emphasis placed on social security, economic activities and other services as reflected in their increased share in public expenditure.

Figure 2.4: Public sector expenditure shares 2000/01 and 2010/11



Source: McIntyre *et al.* (2007) based on (National Treasury, 2005, 2007)

It is important to note that there are indirect effects on health of increasing the share of government expenditure on social welfare and safety nets such as child support grants, old-age pensions, etc. This has been recognised as related to the social determinants of health. On this McIntyre *et al.* (2007 p.7) note that “the spending increases in social security and welfare and in economic services have greatly contributed to shaping healthy livelihoods”.

Based on data from the World Bank (2011b), though South Africa has the largest GDP in sub-Saharan Africa and accounts for about a third of the region’s total GDP, its income inequality is not only one of the highest in the world but is worsening. The Gini index, a measure of income inequality, increased from 0.65 in the late 1990s to 0.72 in 2005/2006 (Statistics South Africa, 2008c). Expressed differently, in 2005/06 the distribution of income in South Africa is such that the poorest 10% of the population shared only about R1.1 billion (representing about 0.1% of total incomes) compared to R381 billion (representing 51%) by the top 10% of the population (Statistics South Africa, 2008c). The increase in income inequality, as reflected in the rise in the value of the Gini index, is associated with *inter alia* inequality in the distribution of skills and a high level of poverty (du Toit, 2002). There are also correspondingly large inequalities in socio-economic status and access to social services between population groups, provinces and socio-economic groupings (Coovadia *et al.*, 2009, Mooney and Gilson, 2009). For example, relatively well-off provinces such as the Western Cape and Gauteng have the lowest poverty rate, highest private health insurance (called medical schemes in South Africa) coverage rate and public health spending per capita, and better access to potable drinking water

when compared to relatively poorer provinces (Coovadia *et al.*, 2009, Development Indicators, 2010). These issues are often considered as social determinants of health as they have direct or indirect impacts on the overall health of the population.

2.3 HEALTH INDICATORS IN SOUTH AFRICA

One of the major indicators of health for any developing country is the number of children who die before celebrating their first birth day (Department of Health *et al.*, 2007). In South Africa, in 2007, the infant mortality rate is estimated at 53 per 1,000 live births. This has remained fairly the same over the last decade (Republic of South Africa, 2010). A summary of some selected health indicators are presented in Table 2.1.

Table 2.1: Selected health indicators, South Africa

Health indicator	Value
Infant mortality rate (/1,000 live births) – 2007	53
Under 5 mortality rate (/ 1,000 live births) – 2008	62
Maternal mortality rate (/100,000 live births) – 2008	410
HIV prevalence (/1,000 adults aged 15-49 years) – 2008	178
HIV/AIDS mortality rate (/100,000 population) – 2009	627
Tuberculosis prevalence (/100,000 population) – 2009	808

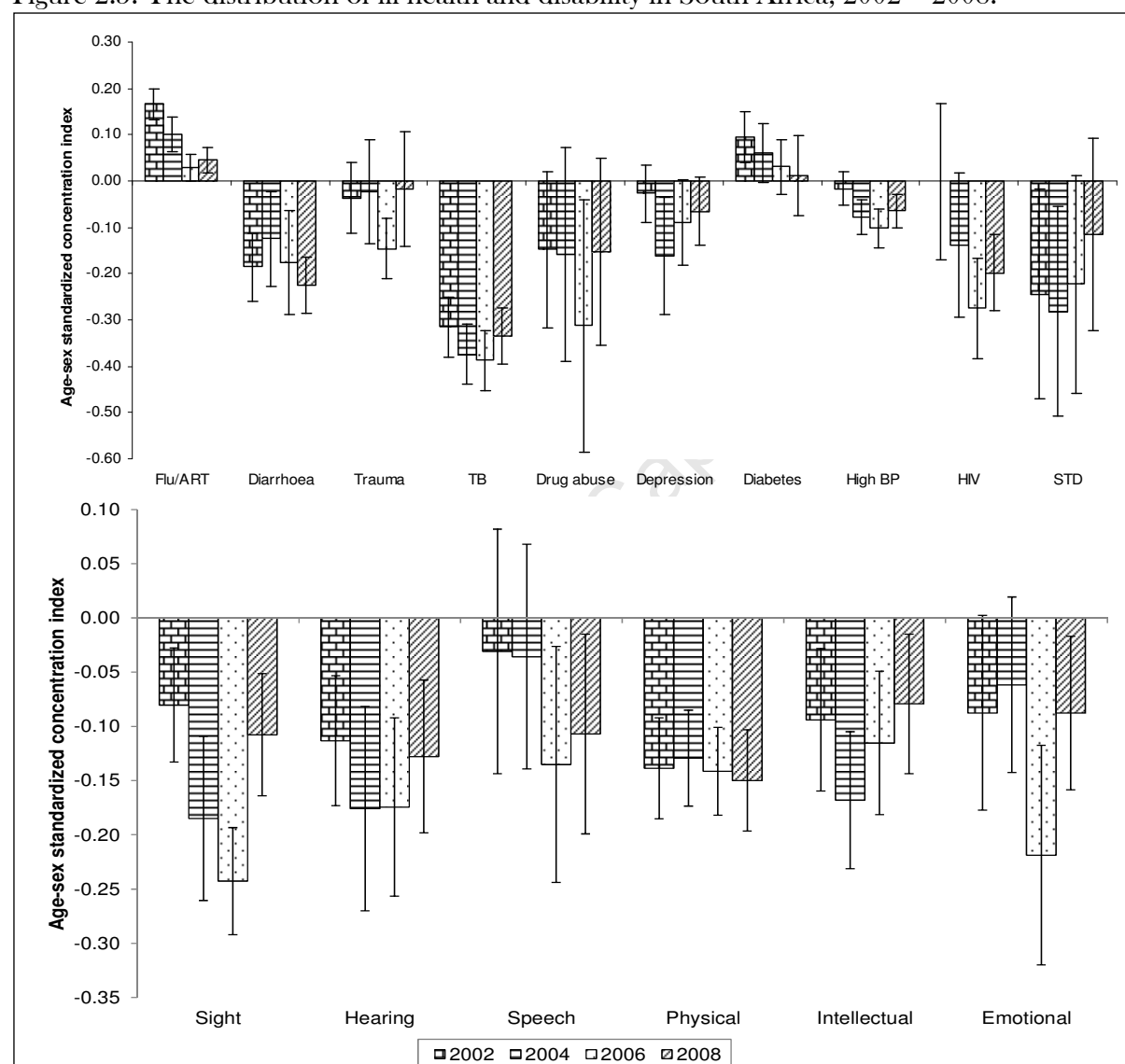
Sources: Republic of South Africa (2010), World Health Organization (2011)

The burden of HIV/AIDS and tuberculosis in the country is very high in comparison to the average burden in sub-Saharan Africa. In 2009 the prevalence of tuberculosis was estimated at over 800 per 100,000 people. Also mortality resulting from HIV/AIDS stands at about 627 per 100,000 people. The high burden of disease in South Africa has been extensively reported in Bradshaw *et al.* (2006). In 2000 the major causes of death included HIV/AIDS, cardiovascular diseases, infectious diseases, and other non-communicable diseases (Bradshaw *et al.*, 2006). The indicators of health presented in Table 2.1 follow the socio-economic gradient (i.e. there is a strong relationship between socio-economic status and health: the poor, relative to the rich, have worse indicators (see McIntyre *et al.*, 2007)).

The burden of the major categories of ill-health and disability, as shown in Figure 2.5, is greater among poorer than richer socio-economic groups. The authors used several rounds of the nationally representative annual General Household Surveys to show that this distribution also

includes non-communicable diseases that are traditionally considered as diseases of affluence¹. They are now being reported by poorer socio-economic groups (Ataguba *et al.*, 2011). Though the prevalence of disability in South Africa is about 5% (see McIntyre *et al.*, 2007), the distribution of this burden across households indicates that they are concentrated more among the poorer socio-economic status groups than acute and chronic illnesses (Ataguba *et al.*, 2011).

Figure 2.5: The distribution of ill-health and disability in South Africa, 2002 – 2008.



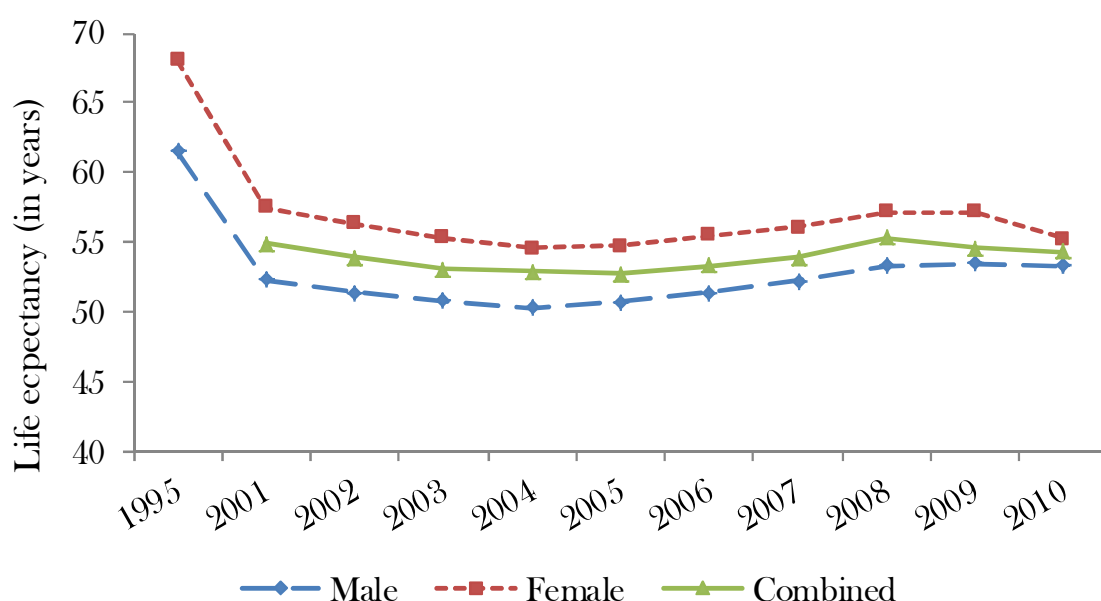
Source: Ataguba *et al.* (2011)

Note: Error bars represent 95% confidence interval; ART = acute respiratory tract infection; TB = tuberculosis; High BP = high blood pressure; HIV = human immunodeficiency virus; STD = sexually transmitted disease.

¹ A positive concentration index signifies that the distribution of ill-health/disability is higher among the richer socio-economic groups while a negative index means that the distribution is higher among the poorer socio-economic groups.

Average life expectancy in South Africa has declined from its 1995 figure. As shown in Figure 2.6 average life expectancy declined from about 65 years in 1995 to 54 in 2010. The gap between the average life expectancy of females and males has narrowed over the years. For instance while the difference between average life expectancy of females and males was 5.2 years in 2001 by 2010 the gap was less than 2 years. The ‘convergence’ can be attributed to the growing burden of HIV/AIDS (Bradshaw *et al.*, 2006) which disproportionately affects females more than males (Development Indicators, 2010, Cleary *et al.*, 2011).

Figure 2.6: Average life expectancy, South Africa, 2001-2010



Source: du Toit (2002), Development Indicators (2010)

Inequities also exist in the distribution of health indicators across provinces and population groups in South Africa. For instance well-resourced provinces have better health indicators than their counterparts. Female average life expectancy (2001-2006) in the richer Western Cape Province is estimated at 60.7 years compared to 48.3 years in Free State Province (Development Indicators, 2010).

2.4 BRIEF OVERVIEW OF THE SOUTH AFRICAN HEALTH SYSTEM

The pre-democratic (pre 1994) period was characterised by the apartheid system that impacted on the South Africa health care system. During this period the health care system was highly

fragmented with different population groups having their own health department and administration (Kale, 1995, Coovadia *et al.*, 2009). Health policy was characterised by the objective of maintaining economic and political power, and a higher quality of life for the minority white population (McIntyre *et al.*, 2007). Health services for the black majority were heavily underfunded and rural areas were neglected (Kale, 1995, McIntyre and Gilson, 2002). Access to health services was highly segregated along the lines of race as there were separate hospitals and other public health facilities for the blacks and the whites (McIntyre *et al.*, 2007).

During the period of democratic transition there were 14 separate health departments operating in the country and the post-apartheid government was faced with huge challenges of redressing existing inequalities and inequities. Following the formal adoption of the 1996 constitution, all the health administrations were amalgamated into one national and nine provincial health departments (McIntyre *et al.*, 2007, Coovadia *et al.*, 2009) with considerable importance attached to primary health care (Republic of South Africa, 2004). In order to redress existing inequalities and inequity in health care access and utilization inherited from the apartheid era, user fees were initially abolished for public health services provided to children less than six years, pregnant and lactating mothers. This initiative was further expanded to cover all public sector primary health care services in 1996 (McIntyre *et al.*, 2007).

The 1996 constitution brought a fiscal federalism structure that gave autonomy to individual provinces in inter-sectoral public resource allocation. The outcome was disparities in terms of priorities that different provinces accord to different sectors (McIntyre and Doherty, 2004). However such disparities have greatly reduced in recent times.

The South African health system, like many health systems in the world, is tiered, i.e. it consists of both private and public sectors. Currently the public sector, with services largely funded through allocations from general tax revenue and a small portion derived from local government revenue and user charges (Gilson and McIntyre, 2007), is structured along the lines of the three tiers of government (national, provincial and local). The respective roles of each tier of government have been spelt out in the 2004 National Health Act (Republic of South Africa, 2004, McIntyre *et al.*, 2007). This is summarised in Table 2.2.

Table 2.2: Health services roles of different spheres of government

Sphere of Government	Responsibilities
National	<ul style="list-style-type: none"> • Formulate national health policy in collaboration with provincial health ministers and officials and local government representatives (via the National Health Council) • Identify national health goals and priorities and monitor implementation progress • Issue guidelines for the implementation of national health policy • Issue and promote adherence to norms and standards (for specific health services and for the training of human resources for health) • Facilitate and promote: health and healthy lifestyles; community participation in health service planning, provision and evaluation; provision of environmental pollution control services; provision of services for communicable and non-communicable diseases • Conduct and facilitate health systems research • Integrate annual plans of the national and provincial health departments • Participate in inter-sectoral and inter-departmental collaboration • Liaise with ministry of health in other countries and with international agencies
<u>Provincial level</u>	
Provincial	<ul style="list-style-type: none"> • Provide specialised hospital services • Provide ambulance and other emergency services, forensic & medico-legal services, occupational health services, environmental pollution control services • Facilitate and promote the provision of comprehensive primary health services and community (district) hospital services • Co-ordinate the funding and financial management, and provide technical and logistical support to districts • Undertake similar activities outlined above, but at the provincial level (e.g. facilitate and promote healthy lifestyles and community participation, participate in inter-sectoral collaboration, etc.) • Plan and manage human resources development and the provincial health information system • Plan, co-ordinate and monitor health services • Plan the development of public and private hospitals and other facilities
<u>District level (District Health Council)</u>	
Local	<ul style="list-style-type: none"> • Promote co-operative governance (between provincial & local government) • Ensure co-ordination of planning, budgeting, provisioning and monitoring of all health services that affect residents of the health district <ul style="list-style-type: none"> • Provide 'municipal health services', including water quality monitoring, food control, waste management, surveillance and prevention of communicable diseases (excluding immunisations), vector control, environmental pollution

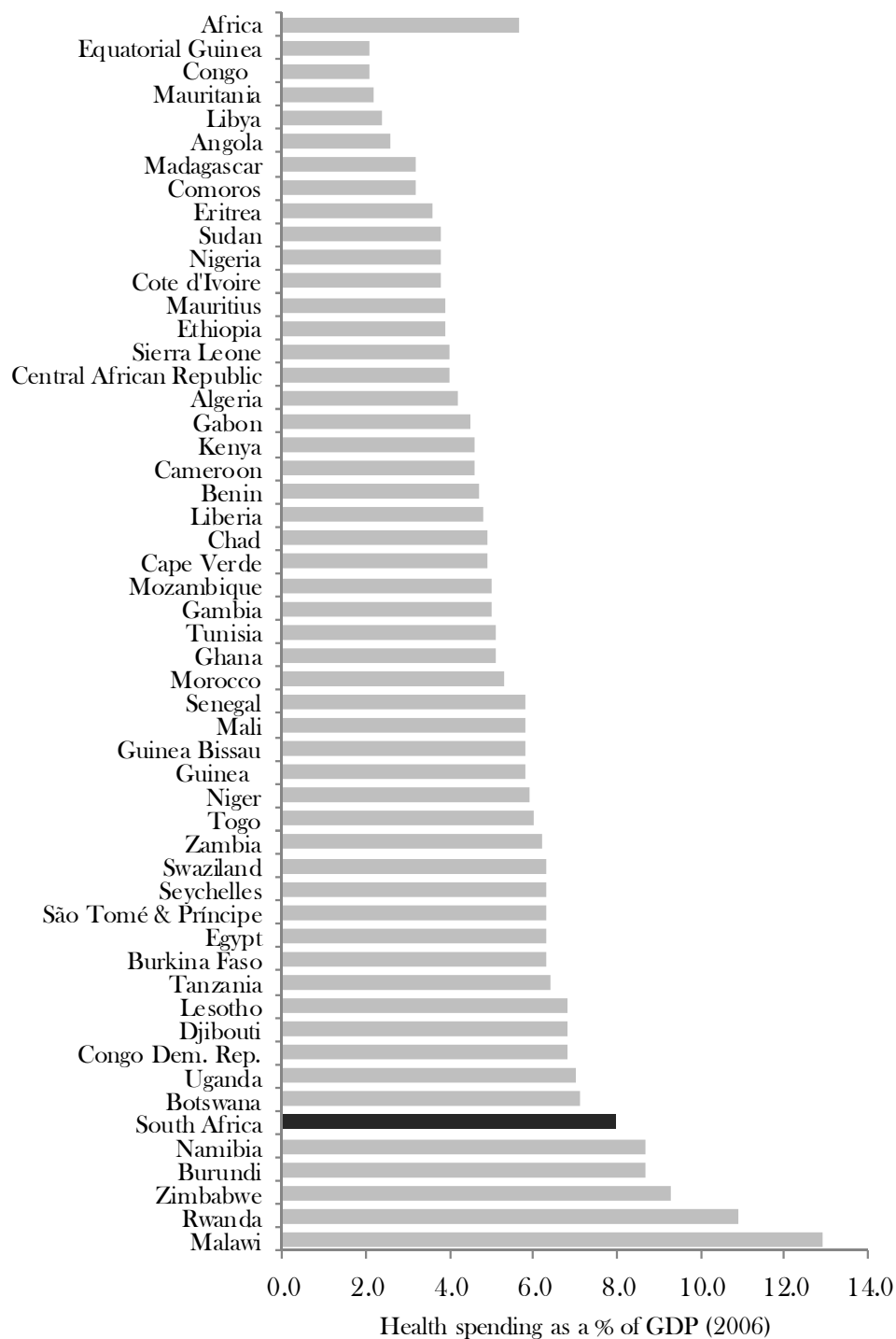
Sphere of Government	Responsibilities
	control, disposal of the dead and chemical safety.
	<ul style="list-style-type: none"> • Primary care services if agreed with the province, on the basis of a service level agreement (specifying services to be rendered, financial resources to be made available by the provincial health department and performance standards)

Source: Republic of South Africa (2004) reproduced in McIntyre *et al.* (2007).

The national department of health is charged with the responsibility of overall guidance, coordination and national health policy development. The provincial departments, with relative autonomy, are responsible for provincial health policy in line with the broad national policy framework. They are the main providers of health services through hospitals and primary care clinics (Gilson and McIntyre, 2007). Local governments are in charge of environmental health services and other non-facility based primary health care services (Gilson and McIntyre, 2007). Similarly the country operates a three-tier public hospital structure (tertiary, regional, and district) and primary health care system (comprising clinics and community health centres) run mainly by nurses (Gilson and McIntyre, 2007, Coovadia *et al.*, 2009).

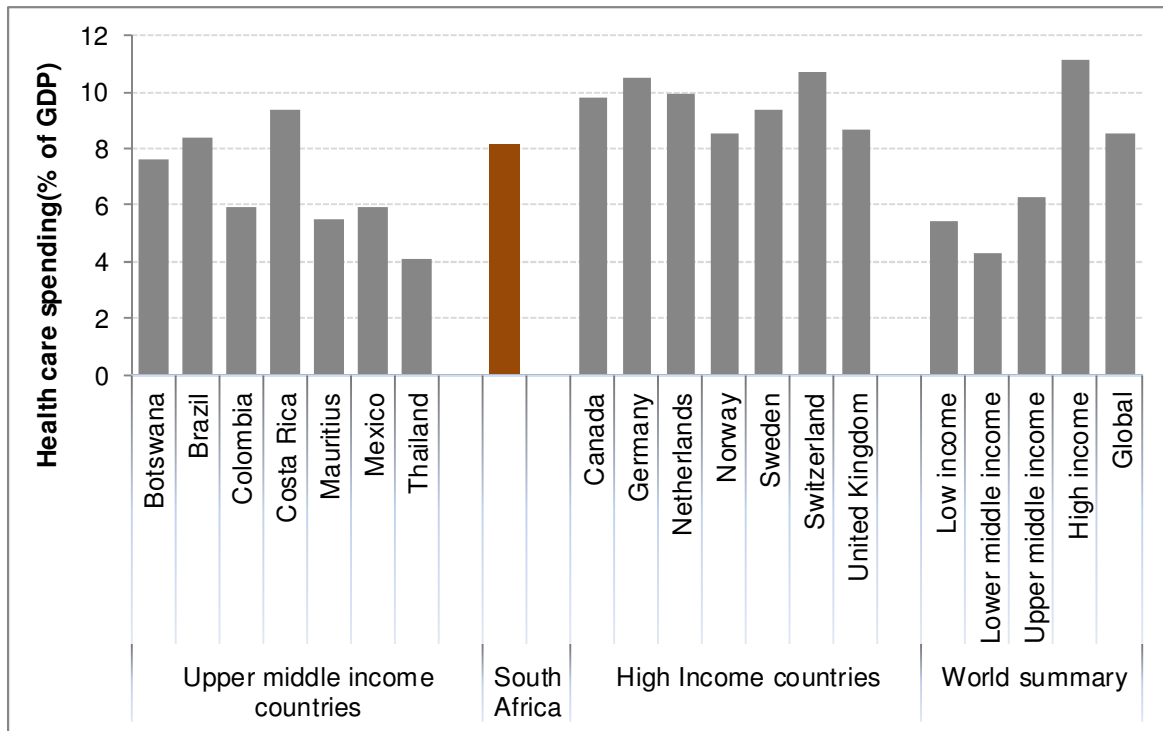
The private health sector on the other hand comprises general practitioners, specialists, pharmacies, private hospitals and traditional health care providers (Gilson and McIntyre, 2007). Though total health care expenditure accounts for over 8% of the country's GDP (see Figure 2.7), the health sector continues to face several equity challenges. In comparison with other African countries, as shown in Figure 2.7, South Africa has one of the largest health care spending in total GDP. The share of health expenditure in GDP also mimics that of some selected high income countries as shown in Figure 2.8.

Figure 2.7: Health care spending as a proportion of a country's Gross Domestic Product (GDP)



Source: African Economic Outlook (2010a)

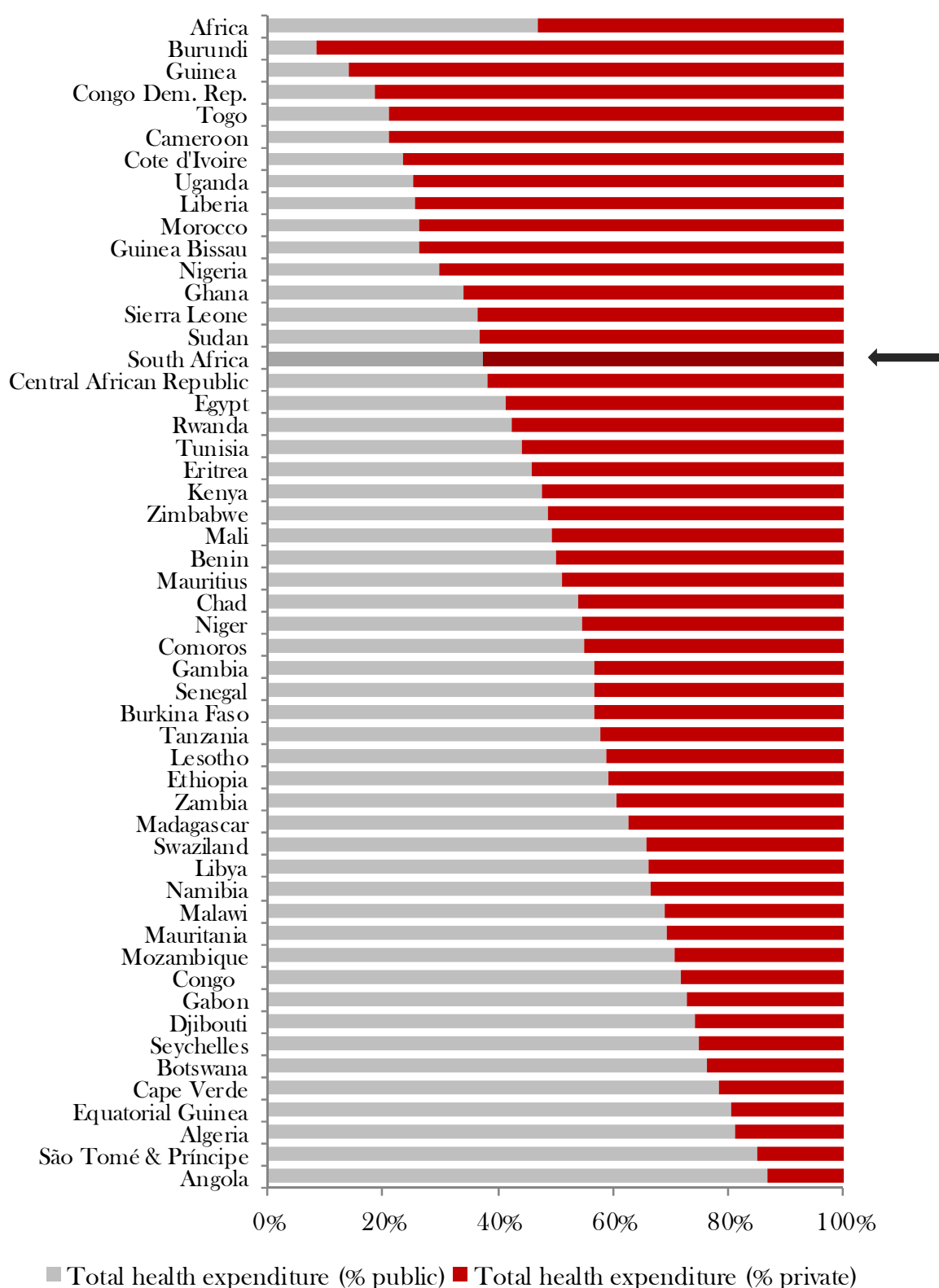
Figure 2.8: Health care spending as a proportion of GDP for selected countries, 2009



Source: World Health Organization (2011)

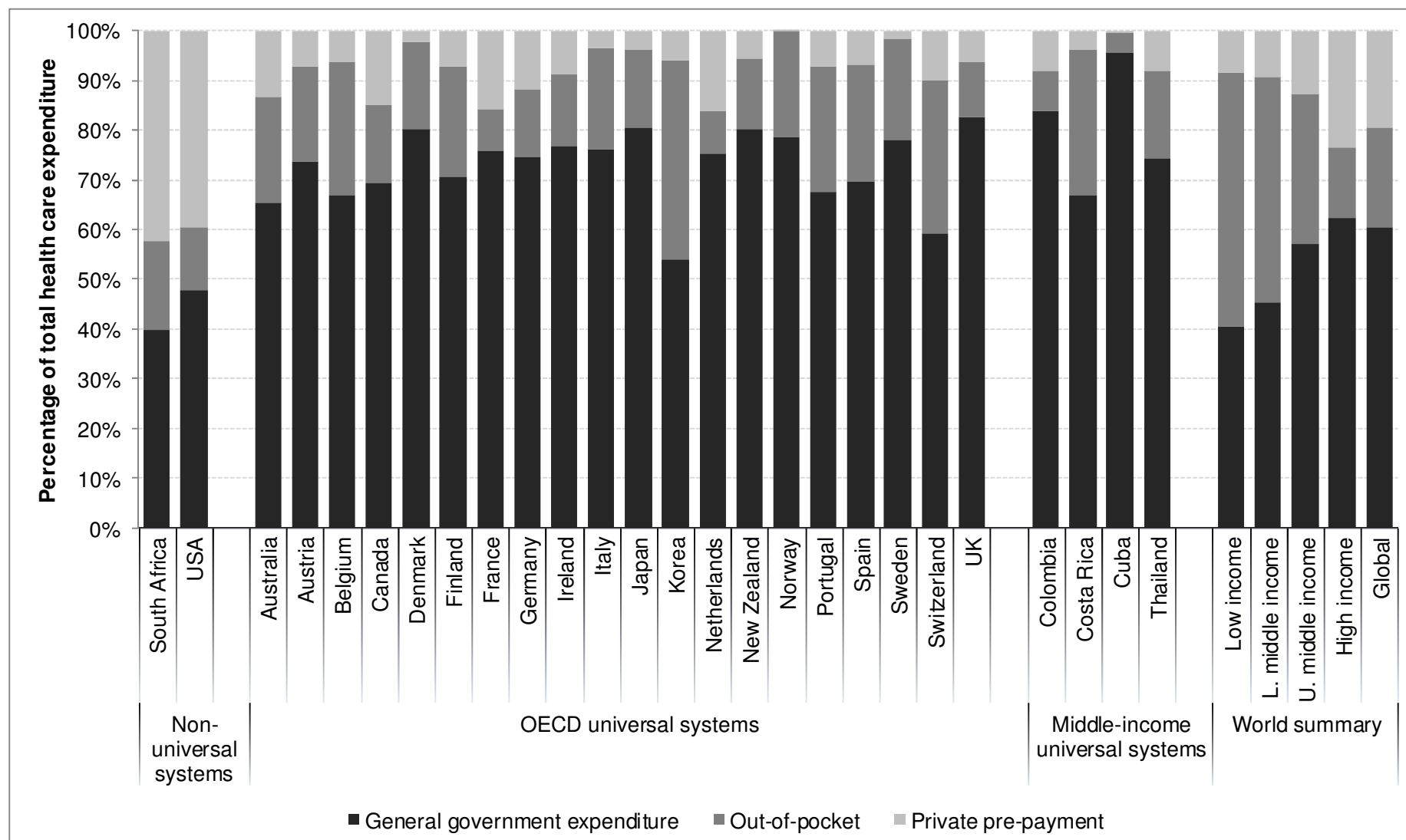
The challenges faced by the South African health system are largely traceable to the legacy of the apartheid era (Kale, 1995). For example over 50% of both financial and human resources are located in the private health sector (Coovadia *et al.*, 2009). In terms of health care expenditure, as shown in Figure 2.9, the private sector alone accounts for over 60% of total health care expenditure. When the private sector's share is disaggregated as shown in Figure 2.10, private pre-payment in the form of voluntary private insurance dominates. Data from the World Health Statistics indicate that South Africa has the highest share of private pre-payment in total health care financing in the world (World Health Organization, 2011). As also shown in Figure 2.10, the share of public health care expenditure in total health care expenditure in South Africa is low in comparison to the USA that also has a high share of private pre-payment in total health care financing and without a universal health system. In fact this share is lower than the average shares for low income, lower middle income, upper middle income and high income countries (World Health Organization, 2011).

Figure 2.9: Health care expenditure (public vs. private) in Africa, 2006



Source: African Economic Outlook (2010a)

Figure 2.10: Health care expenditure (public vs. private) in selected countries, 2009



Source: World Health Organization (2011) based on McIntyre (2012).

In South Africa it is estimated that over two-thirds of private hospitals are located within three of the nine South African provinces (Coovadia *et al.*, 2009). Inequity also exists within and across public facility levels as the government devotes a large portion of funding to tertiary and academic hospitals to the detriment of non-hospital primary care services that cater mainly for the poor (Coovadia *et al.*, 2009). These challenges further require that available resources be used in an efficient manner and that resource allocation decisions are based on evidence.

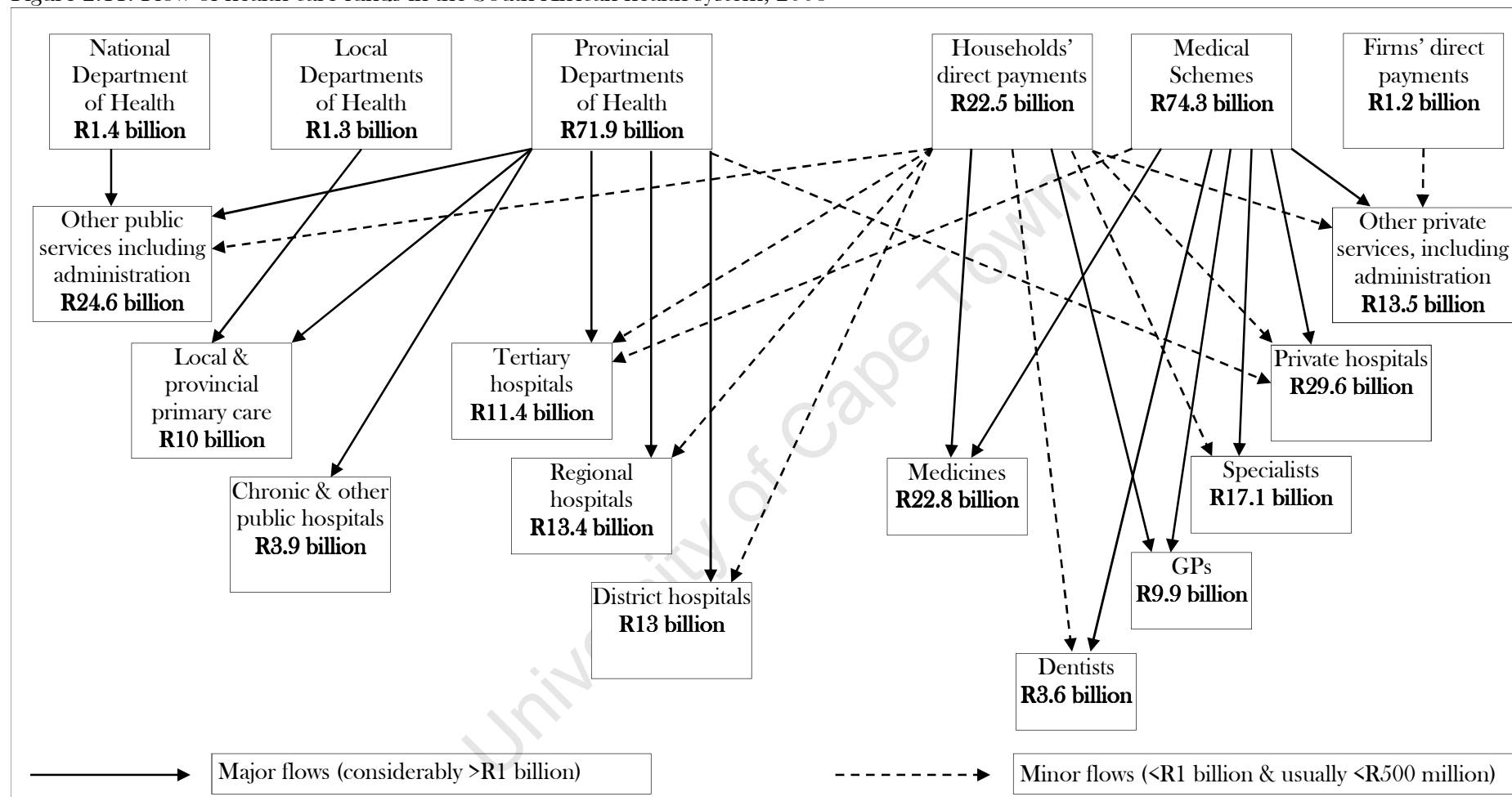
2.5 HEALTH CARE EXPENDITURE AND FINANCING IN SOUTH AFRICA

Internationally, it has been acknowledged that the way in which 'health systems are financed largely determines whether people can obtain needed health care and whether they suffer financial hardship as a result of obtaining care' (Carrin *et al.*, 2007 p.652). In South Africa, health care is financed through both public and private mechanisms. Public mechanisms mainly comprise general taxes while private mechanisms consist of direct household out-of-pocket payments and contributions to medical schemes and health insurance. Between the period 2005/06 and 2008/09, nominal total health care spending increased annually by 16.7% (National Treasury, 2009b). In 2005 total health care expenditure was about 7.7% of GDP and it totalled R 108 billion (McIntyre *et al.*, 2007). By 2008 about R 173 billion (i.e. about 7.6% of the country's GDP) was raised to finance health care (see McIntyre *et al.*, 2012). While there have been sustained nominal increases, the share of health care financing in GDP declined from about 8.8% in 1999/2000 (McIntyre and Doherty, 2004) to 7.6% in 2008. This is especially the case for government financing that has not been keeping pace with population growth and inflation (McIntyre *et al.*, 2012).

The 1996 Growth, Employment and Redistribution macroeconomic policy contributed to the fiscal contractions that affected public health care in the country. The limits set on the country's tax-to-GDP ratio limited government's spending power. This affected government spending on many social service sectors including health care (McIntyre and Doherty, 2004). Though the trend is currently being reversed (McIntyre *et al.*, 2012), the impact created by previous deprivation is still evident in the health sector. This is even compounded by the current burden of HIV in the country (Arndt and Lewis, 2000).

A description of the flow of health care financing in the South African health system is summarised in Figure 2.11. The Figure shows how funds flow from different sources (public or private) to health care providers. In 2008 it is estimated that about 43% of total health care funds is accounted for by general government revenue. This is almost equal to that accounted for by private medical schemes alone. Because of the size of the central or national government (accounting for >90% of government revenue), it is able to exercise greater leverage on how health care funds are spent in South Africa (McIntyre *et al.*, 2012). The trend depicted in Figure 2.11 has remain fairly similar over the years (see McIntyre and Doherty, 2004, McIntyre *et al.*, 2007). The private financing sources (i.e. household out-of-pocket payments and medical schemes) together account for about 57% of total health care resources in South Africa. Though private medical schemes account for about 43% of total health care funds, they cover less than 16% of the population that comprises mainly the rich.

Figure 2.11: Flow of health care funds in the South African health system, 2008

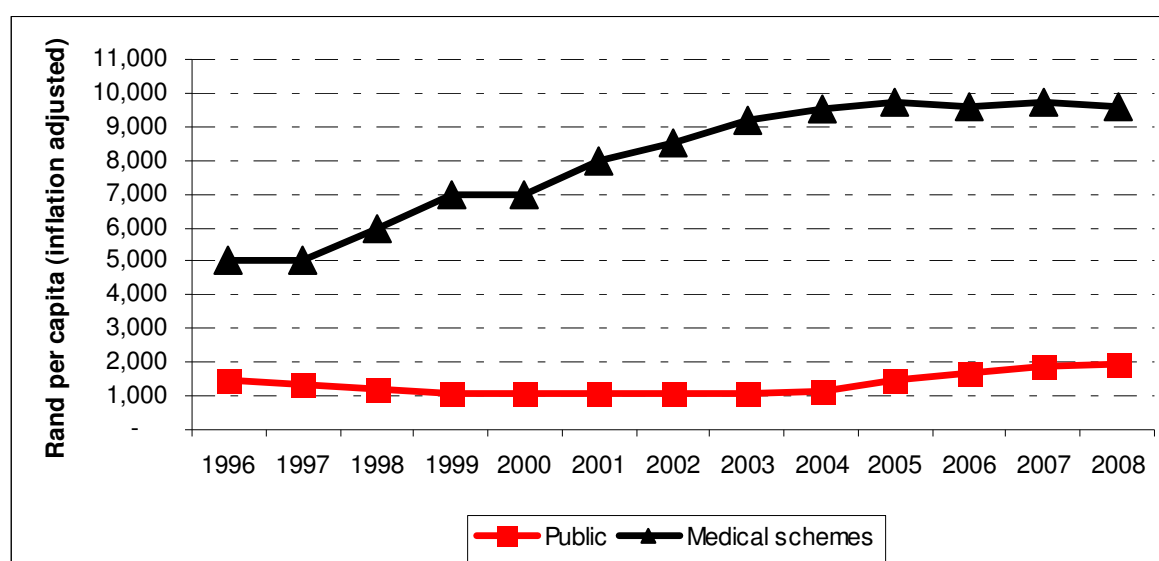


Source: See McIntyre *et al.* (2012).

Note: US\$ 1 ≈ R 7 (South African Rand)

As shown in Figure 2.12, in 2008 and in real terms about R10,000 was spent per medical scheme member while only about R1,900 was spent per individual dependent on the public sector. The “maldistribution of resources between the public and private health sectors, relative to the population that each serves, reflects inefficiencies and inequities that contribute to South Africa falling far short of the Millennium Development Goals” (McIntyre *et al.*, 2009 p. 725).

Figure 2.12: Trends in real per capita spending by medical schemes and the public health sector



Note: Medical schemes figures are per scheme member while public health sector figures are based on those dependent on the sector.

Source: McIntyre *et al.* (2007) based on data from Council for Medical Schemes Annual Reports (for medical schemes); National Treasury annual Budget Reviews (for public spending); Statistics South Africa (for CPI and population).

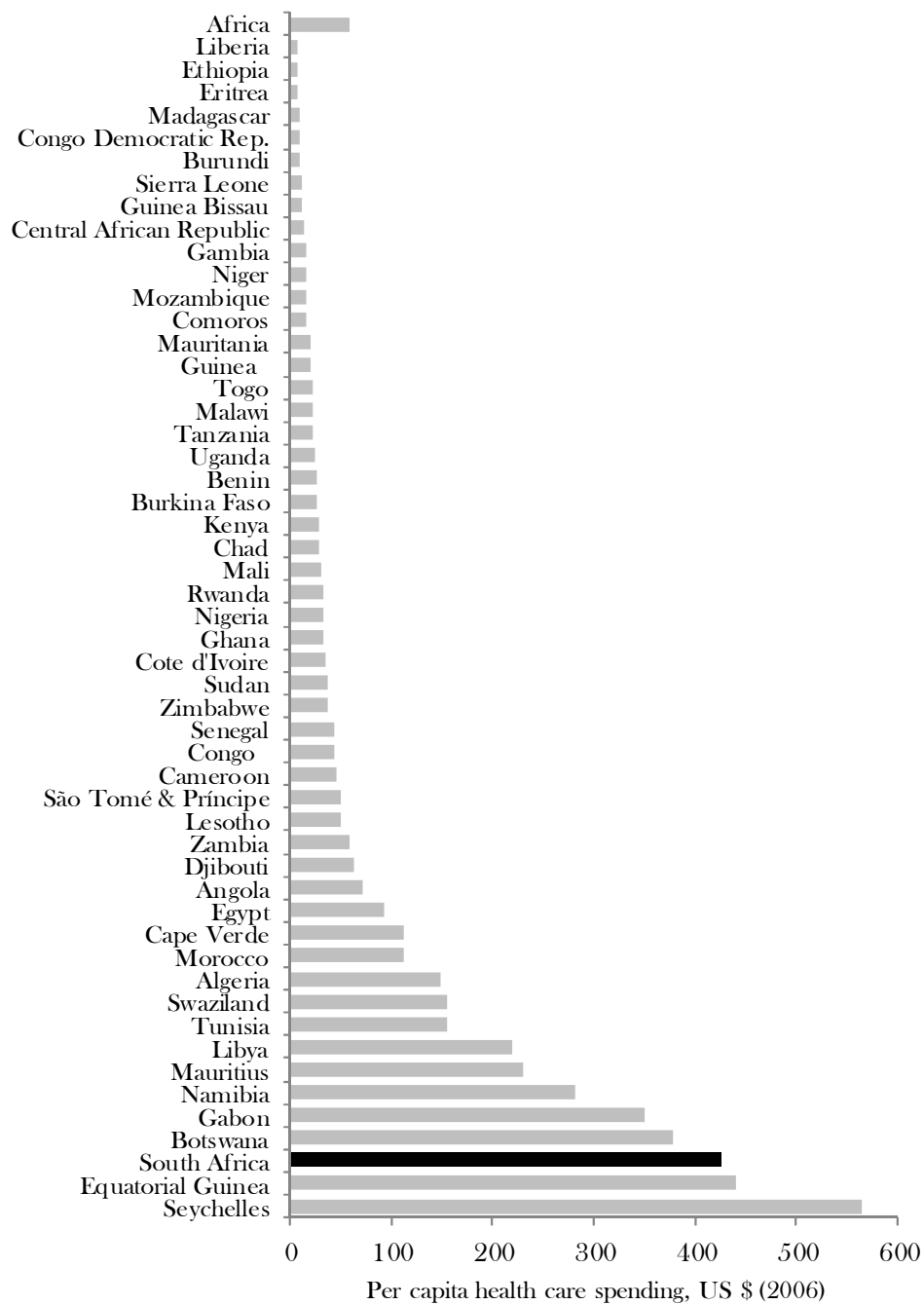
Based on Figure 2.12, the gap between per capita spending on medical scheme members and public sector spending was lower in the late 1990s than in the late 2000s. In 1996 for instance, medical scheme spending per capita was about triple that of public spending and by 2004, the gap had grown to more than seven times, and has remained at a similar level since then (see McIntyre *et al.*, 2012). It was only at the beginning of 2003 that real public spending per capita began to rise, but these increases were relatively small in comparison to the pace of increases seen in private medical schemes’ spending per capita over the decade. On the whole, combining medical schemes’ spending, public spending as well as private out-of-pocket spending, per capita spending on health, in real terms, has been increasing in South Africa. While this may seem positive overall, as shown in Figure 2.12, it is largely attributed to increased medical schemes’ expenditure and contributions, which is only benefiting the minority who are members of these schemes. This still creates challenges for the health sector.

2.6 HEALTH CARE FINANCING CHALLENGES IN SOUTH AFRICA

Internationally, the World Health Organisation has noted, based on the experience of high income countries, that “the tendency of many governments of middle-income countries to attempt to shift the finance of clinical health services to the private sector, especially via fee-for-service payments, runs the risk of dramatic cost escalation and would virtually ensure that a substantial proportion of the population will lack financial access to services in time of need.” In lieu of this, they have encouraged countries to mobilise needed public finance (World Health Organization, 2001 p.62).

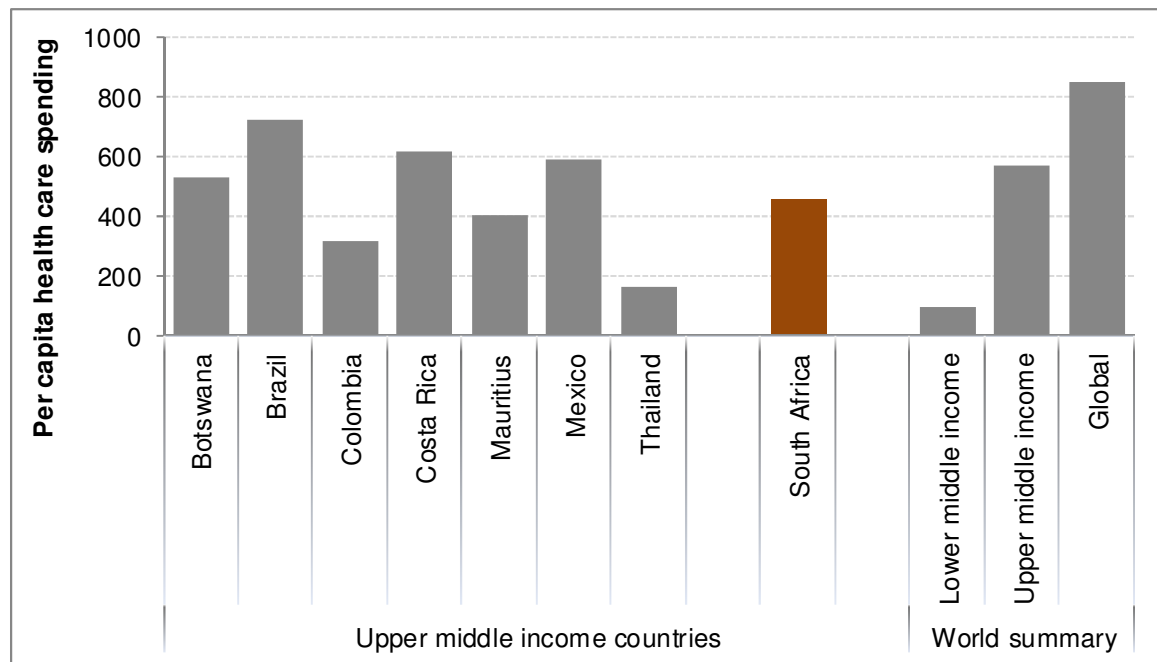
As demonstrated in Figure 2.13, South Africa has one of the highest total per capita health care spending levels (i.e. from public and private sources) in Africa. As shown in Figure 2.14 per capita health care spending in South Africa is also higher than those of other middle income countries (such as Colombia and Thailand) that have universal systems. This high per capita health care spending coupled with the information contained in Figure 2.12 implies that there is relatively adequate per capita health care spending to provide more than the basic level of care for everyone in South Africa. This follows from the Commission on Macroeconomics and Health’s estimates of health care resource requirements (World Health Organization, 2001). They provide an average estimate of US\$ 237 for 2007 and US\$ 259 by 2015 as the required per capita health care spending for scaling up priority health interventions in middle income sub-Saharan African countries (World Health Organization, 2001). In 2006, as shown in Figure 2.13, per capita health care spending in South Africa was US\$ 425. This amount clearly exceeds the level estimated by the World Health Organisation for 2007 and 2015. Even at this high level of per capita health care spending, it has been shown that the use of health services (private and higher level public services) and the benefits from using these services, measured in monetary terms, are skewed toward the rich (Ataguba and McIntyre, 2012). South Africa has also been described as a classic case of where the inverse care law (Hart, 1971) persists. This is a case where those with the greatest need for health care services use and benefit less than those with relatively less need for health care services.

Figure 2.13: Total expenditure on health per capita in Africa, 2006 (in US\$)



Source: Data from African Economic Outlook (2010a)

Figure 2.14: Health care spending per capita for selected upper middle income countries, 2009



Source: World Health Organization (2011)

An important question then has been: why do the majority of South Africans still experience problems with accessing quality health services given that there is relatively abundant financial resources for health care? (McIntyre *et al.*, 2012). While the answer to this question is not straightforward, the problem lies in the way health care is financed and the way such funds are used in purchasing health services (World Health Organization, 2000). The absence of an integrated pool of resources and inequity in financing health care could contribute to the limited access to health care (McIntyre *et al.*, 2012).

In South Africa, it can be deduced that one of the most significant challenges facing the health sector is the distribution of resources between public and private sectors relative to the population that each serves (McIntyre *et al.*, 2007). The private sector that serves a minority population accounts for over half of total health care funds in comparison to the public sector. In addition to this challenge, another major issue is the level of fragmentation of resources that exists within the private sector. Currently, there is a great degree of fragmentation *within* the medical schemes' market. As at January 2010 about 105 medical schemes were registered in South Africa (Council for Medical Schemes, 2010). These schemes have different population pools, resulting in a general lack of cross-subsidization in the overall medical schemes environment (McIntyre *et al.*, 2008). While there were moves towards developing a risk-

equalization mechanism across the current medical schemes, this has never been implemented (McIntyre and Doherty, 2004, McIntyre *et al.*, 2008).

McIntyre *et al.* (2008 p.873) have argued that “[t]he key pooled funding mechanisms for health care are tax (and donor) funding and health insurance”. In terms of tax funding in South Africa, the health sector share of the government budget has been declining from about 11.5% of the total government budget in 2000/2001 to about 10.9% in 2007/2008, which is contrary to the Abuja declaration committed to by African heads of States in 2001. The commitment requires countries to allocate up to 15% of total government budget to the health sector (McIntyre *et al.*, 2008). These challenges have resulted in the need for urgent health care financing reform in South Africa.

2.7 CURRENT HEALTH CARE FINANCING REFORM IN SOUTH AFRICA

Since the 1940s, South Africa has considered a series of proposals for health care financing reforms and none have been fully actualised to date. Lack of political support, the obstructive role of some stakeholders and the lack of general buy-in has contributed to the non-implementation. In 2007 in Polokwane, the African National Congress (ANC) – the South African ruling party – committed itself to the establishment of national health insurance (NHI). This is due largely to concerns about the challenges facing the South African health system (within both the public and private sectors). It reflects growing concerns for the poor who sometimes cannot utilise health services due to high costs (not only of health services but transport to access services), employees complaining about the escalating contributions to medical schemes, as well as failed attempts in the past to introduce health care financing reforms (McIntyre *et al.*, 2009).

The NHI proposal by the ANC has generated considerable debate in the health policy arena, particularly among key health sector stakeholders and academics. Recently a Green Paper has been released that provides an overview of the proposed NHI (Department of Health, 2011). Based on the recently released Green Paper, the NHI is expected to be fully operational by 2025. The process is expected to comprise three phases. During these periods, structures are expected to be put in place to handle the anticipated increment in health service utilization rates. The central theme underpinning the proposed NHI is that the South African health

system needs to be restructured to ensure better performance and that it meets the needs of the people whom it is serving.

The NHI is proposed to have its funds pooled into a single autonomous public entity called the National Health Insurance Fund (NHIF). This pool of funds will be used to purchase quality health services on behalf of the population (Department of Health, 2011). It is expected that, in addition to public providers, accredited private providers will provide services under the NHI. While the exact nature of this accreditation and purchasing mechanism is expected around 2013/14 (Department of Health, 2011), it is seen as critical to address the skewed distribution of health care resources (human resources and financial) that is in favour of the private sector (see McIntyre *et al.*, 2009). Similarly, the proposed NHI does not preclude the existence of private medical schemes. However, enrolment in these schemes is optional for those who wish to obtain additional health care cover, but must be paid for over and above any mandatory NHI contribution (Department of Health, 2011).

The major aim of the proposed NHI is to achieve universal coverage (i.e. “provide universal financial protection against the costs of using health services when needed” (McIntyre *et al.*, 2009)) such that even those who cannot afford to pay for health care at all or at the point of utilization, will be able to use quality health care services without the fear of financial risks and other associated losses.

Part of the debates about the proposed NHI relate to concerns raised about the feasibility and the likelihood of success of this large-scale reform of the health system. The NHI seeks to make the public sector ‘a provider of choice’, both because it is the major provider of health care already and because there is a greater possibility of cost containment within publicly (than privately) provided health services if effectively managed. However, there are currently serious challenges and inadequacies in the public health sector (see Coovadia *et al.*, 2009, McIntyre *et al.*, 2009 for example). The legitimate concerns about the lack of preparedness of the public health sector are not in essence a valid reason for not pursuing an NHI; instead, they relate more to the pace at which an NHI can be implemented. The government has taken this into consideration in the design of the NHI (Department of Health, 2011). While many point exclusively to deficiencies in the public sector, it is important to acknowledge that the private sector has its own share in inefficiencies and perverse incentives associated with fee-for-service payments for those who are privately insured. In particular, medical schemes’ contribution levels are becoming increasingly unaffordable and they do not provide good financial

protection as members still have to make substantial co-payments (see also McIntyre *et al.*, 2009, McIntyre *et al.*, 2012). Another issue related to affordability and the pace of introducing the proposed NHI is the need to recognise the impact of the recent economic recession, following the global financial meltdown, on the feasibility of having expanded fiscal resources for the NHI in the short-term. The growing burden of disease in South Africa, particularly in relation to the HIV/AIDS pandemic, is also cited as a potential source of escalating the resource requirements for adequately covering the whole population. However, the issue here is that of priority setting and a coherent plan taking into account all the competing needs for resources. Recent estimates have shown that under the NHI total health care spending as a proportion of GDP could be lower than or slightly higher than what it currently is (McIntyre, 2010c, Department of Health, 2011). Though estimates that are almost equivalent to the current proportion have been produced (McIntyre, 2010c, McIntyre and Ataguba, 2012), the estimates contained in the policy paper indicates that NHI in the long-run will cost about 6.2% in comparison to the current 8.5% of total health care spending in GDP (Department of Health, 2011). The lower estimate contained in the Green Paper, however, excludes contributions to private health insurance by those who will continue to purchase insurance as top-up cover.

2.8 SUMMARY

South Africa has gone through a phase of transition from the apartheid era to a new democratic country. However the country still faces some challenges. These challenges include those that are related to broader economy-wide issues such as redistribution, poverty, unemployment and economic stabilisation; and those that are peculiar to the health sector including the health care financing system. Currently income inequality and inequalities in access to social services are huge and are compounded by the skewed distribution of disease burden including HIV/AIDS. Among the challenges that have been identified for the health care financing system is the fragmentation in the medical schemes market and the absence of a universal system where access to and use of health services is not dependent on income but on the need for care.

Chapter 3

THEORETICAL FRAMEWORK AND LITERATURE REVIEW

3.1 INTRODUCTION

This chapter reviews the literature on equity in relation to health care financing. This review shall cover mainly theoretical debates. It provides a theoretical background upon which subsequent chapters are built. The chapter introduces and discusses the major theories of justice. These include utilitarianism, egalitarianism, libertarianism and Rawlsian postulations. While these theories were not originally based on equity in health care financing, an attempt was made to put the discussions around them within the context of equity in health care financing. They are used to also shed more lights on our understanding of equity in health care financing.

3.2 THEORETICAL REVIEW

“Justice is blind, and fairness requires anonymous rules of arbitration”
(Moulin, 2003 p.1)

What is equity or an equitable distribution? How should the burden of health care financing, like tax, be distributed across individuals, households or groups of people? What principle(s) should be used in distributing this burden? What is the preferred relationship between the health care financing burden and households’ ability to pay? How do we assess if a distribution is fair and/or just? These are some of the questions that philosophers and economists are faced with and have been trying to understand and provide answers to. Different theories provide different ways of addressing the problem and there seems to be no consensus as to what a fair distribution might look like. While the distinction may sometimes be made between fairness and justice (Rawls, 1999), these issues are ones usually discussed under theories of justice. These types of questions are not restricted only to the distribution of the

burden of health care financing. They also apply to any *distribuendum*² of interest. They have their roots in the economic and political philosophy literature (Moulin, 2003) that dates back centuries.

To begin to understand these issues, we need to set the stage for what will follow. The discussion here will be based mainly on explanations that relate to how the burden of health care financing is distributed across different people in a population. A simple analogy is presented here. This will enable us to understand the arguments related to the philosophical discussions on how health care financing burden should be distributed. This simple analogy is based on that presented by Julian Le Grand (1991 p.64-65). Suppose we could observe three individuals³ - Akoji, Unekwu and Ojochide who are seeking health care at a hospital. They are all required to pay for treatment. Assuming that we have no further information about them (e.g. their background, household size, income, habits, etc.) except that they all, for simplicity, present with *the same condition*⁴, and further that Akoji and Unekwu pay the same amounts for treatment, but Ojochide pays an amount that is three times that of each of the others. With only this information, we might be tempted to judge the situation to be *inequitable*. Thus if we define equity as equality of health care payments for equal health condition, then the distribution is inequitable. Suppose further that we now have information on their income levels. If Ojochide is three times as rich as Akoji and Unekwu, we might judge the situation to be equitable, if we define payments being proportional income. Yet again if we define equity in regard to health care payments as a situation where higher income individuals pay a higher share of their income, we would conclude that the burden is not equitably distributed. To be equitable now would require that Ojochide pay more as a proportion of income.

3.3 ECONOMICS AND THE CONCEPTS OF EQUITY AND FAIRNESS

“[w]e must preserve the integrity of the economics profession by not presuming to know what is fair and what is not.” (Slemrod, 1996 p.8)

² These could include opportunities, outcomes, resources or welfare (Roemer, 1996).

³ Here individuals are used. They could be generalised to types or categories of individuals.

⁴ Though it may not be simple to generalise that the three persons here have the same level of illness, this restriction is important in order for us to proceed.

The simple example using Akoji, Unekwu and Ojochide shows that the understanding of an equitable distribution of the burden of health care financing, just as in the case of taxation, including how it should be defined and measured, is not straightforward and no consensus among economists and philosophers exists (Wagstaff and van Doorslaer, 1992, Musgrave, 1996). It is value-laden (Le Grand, 1987a, b, Slemrod and Bakija, 2004) and as suggested by McLachlan and Maynard (1982) “equity, like beauty, is in the mind of the beholder” (quoted in Wagstaff and van Doorslaer, 1992 p.363). This lack of consensus in defining equity has also been echoed by several economists such as Le Grand (1987a p.258) noting that “[e]quity is widely regarded as a term capable of an almost indefinite number of interpretations, dependent solely upon the values of the person using it at the time. This can be contrasted with the situation with respect to, say, efficiency, where something close to a consensus interpretation does exist (that of Pareto)” (see also Le Grand, 1987b). All of this is reflected in the many theories of justice, including their variants, that have been proposed (see for example Foley, 1967, Rawls, 1971, Nozick, 1973, 1974, Dworkin, 1981b). Because there still exist differences among theorists as to how fairness and justice should be defined, in respect to economics, Musgrave writes that “[t]hrough economics cannot establish the ‘correct’ equity norm, it can explore the implications of alternative norms, implications that must be weighted in the balance if an intelligent choice among them is to be made” (Musgrave, 1996 p.342). This advice guides the exposition of the theories of justice that follow and it requires some understanding on how each of the theories can be applicable to various issues of distribution, particularly related to taxes and health care payments, in economics.

The various theories of justice and fairness can be traced to the principle of the *Nicomachean Ethics* by Aristotle⁵ (Moulin, 2003). This philosophy based on ethics, is noted as one of the origins of economics (Sen, 1988). Sen links the other origin to *engineering*. Though Sen would argue against divorcing economics and philosophy (Sen, 1988), it is the case that the “ethical approach has rather substantially weakened as modern economics has evolved” (Sen, 1988 p.7). In fact “[t]he distancing of economics from ethics has impoverished welfare economics, and also weakened the basis of a good deal of descriptive and predictive economics” (Sen, 1988 p.78). While there are no doubts in economics as to the relevance of ethics to welfare economics (Sen, 1988, Konow, 2003), it is important to note that ethical

⁵ This basically requires that *equals* be treated *equally* and *unequals* be treated *unequally*, in proportion to the *relevant* similarities and differences they share (see Moulin, 2003).

considerations more generally shape human behaviour and judgements (Sen, 1988). However some economists still argue that the discourse of equity should be left to philosophers, political scientists, and sociologists (Konow, 2003).

It is possible to point to the advances in economics that can be attributed to the principle of Pareto optimality and to the general theories of welfare economics more broadly. While their application has produced remarkable results in economics, especially as it related to contracts, trade and general equilibrium analysis, Pareto optimality could coexist with an *unfair* distribution (Sen, 1988). This is because for example, “[a] state can be Pareto optimal with some people in extreme misery and others rolling in luxury, so long as the miserable cannot be made better off without cutting into the luxury of the rich” (Sen, 1988 p.32). In fact “[t]he Pareto principle can scarcely be extended to cover judgements of distribution without actually making interpersonally comparative distributional judgement” (Sen, 1988 p.33). This statement from Sen brings in some principles related to equity (Konow, 2003). This is because the objective of economics “is to understand, explain and predict human behaviour in a way such that economic relationships can be fruitfully studied and used for descriptive, prognosis and policy” (Sen, 1988 p. 79). The Pareto optimality principle is closely related to the utilitarian logic⁶ in welfare economics which favours maximization of utilities and social welfare functions. In general while “welfarism is the view that the only things of intrinsic value for ethical calculation and evaluation of states of affairs are individual utilities” (Sen, 1988 p.40), it is argued here that welfarism be enriched by drawing on some ethical principles that go beyond maximization and its consequentialist orientation.

In modern economics, the relevance of *fairness* and *equity* in economic analysis has been stressed and these constructs are increasingly being applied (Kahneman *et al.*, 1986). It is the case that “[j]ustice arguments are now widely invoked to improve theoretical and empirical analysis in nearly every field of economics” (Konow, 2003 p.1188). Areas of application include game theory and the notion of “fairness equilibrium” (Rabin, 1993, Nelson Jr, 2001); the theory of consumer choice (Thaler, 1980); production and supply (Kahneman *et al.*, 1986); health economics (Wagstaff and van Doorslaer, 2000, Mooney, 2009); public and labour economics (Fuchs *et al.*, 1998); poverty, inequality and distributional issues (Atkinson, 1970,

⁶ In the next section a detailed exposition of the utilitarian principle of justice is presented.

Sen, 1976, Anand, 1977, Atkinson, 1987, Duclos and Araar, 2006); regulation economics; social choice (Sen, 1977) and so on (see Konow, 2003).

Moulin (see Moulin, 2003) sets out four basic ideas or principles upon which distributive fairness and/or justice can be interpreted. These are discussed in relation to *relevant similarities or differences* in the context of the *Nicomachean ethics*. They include *compensation, reward, exogenous rights*, and *fitness*. These principles may not be regarded as mutually exclusive (Moulin, 2003). To understand them we draw on the canonical story of the flute and four children that dates back to the time of Philosopher Plato (Moulin, 2003). The flute must be given to one of the four children on the basis of different criteria that can be justified on different grounds. If the first child has fewer toys than the others, and the other three attest to this, then by the compensation principle, this child should be given the flute to compensate her for the differences in the amount of toys she has. If the second child had worked hard and spent several months in making the flute, based on the reward principle, she should have the flute as a reward. If the flute originally belongs to the third child's guardian and the guardian no longer cares to have it, then the child has the right to claim it under the exogenous right principle. If the fourth child is a flautist, and he plays better, and all the others acknowledge this, then the flute should be given to him based on the fitness principle because of the efficiency with which it can be used (Moulin, 2003)⁷.

These principles can be extended to the discourse of equity in health care financing. The principle of compensation in this context is about equalising differences in socio-economic disadvantages. This is described as *equality ex-post* (Moulin, 2003) and requires some sort of egalitarian treatment. It may include granting tax-breaks and structuring payments in such a

⁷ A similar narration of the story is contained in Sen's book - *The Idea of Justice* (Sen, 2009 pp.12-15). Sen however narrates the story using three children. In his narration he shows that theorists with different persuasions such as the utilitarians, economic egalitarians and libertarians may relatively resolve the problem of allocation easily and differently. The egalitarians would favour the child with fewer toys; the utilitarians will favour the flutist; while the libertarians will give the flute to the maker. From these reasoning, we can link Moulin's principles to the theories of justice that follow in the next section. Basically, following Sen's presentation (Sen, 2009), Moulin's principle of compensation is akin to the egalitarian philosophy; the fitness principle is related to the utilitarian philosophy; while the reward principle is linked to the libertarian ideology (see also Fleurbaey, 2001). The exogenous rights principle can be linked to Nozick's formulation of entitlements and holdings (Nozick, 1973). It is also important to note, as Moulin argues that each of these principles may sometimes over-lap and we may not be able to put them into tight compartments.

way as to equalise post-payment incomes. The reward principle could be positive or negative (a penalty). Individuals with poorer health might be assumed to lead unhealthy lives (such as not engaging in exercise, poor eating habits, etc.) and as a result they need more health care. They might be required to pay more as a negative reward or penalty for their unhealthy lifestyle (but noting that ill-health cannot always be attributed to leading unhealthy lives). In the same way, those who lead healthy lives and are healthier should be rewarded by paying less for health care than their counterparts. Exogenous rights can be divided into two – equal exogenous (or *equality ex-ante*) and unequal exogenous rights (Moulin, 2003). Based on *equality ex-ante*, the rights are equally distributed across everyone. Therefore the burden of health care financing should be distributed equally among everybody irrespective of, for instance, who is richer or poorer. The distribution of this burden is considered to be independent of how it affects the welfare of anybody. This is in contrast with the *equality ex-post* implied by the compensation philosophy. If however we note that being born into a particular household (rich or poor) is a matter of an accident of birth, then unequal exogenous rights demand that these inequalities be taken into account in the distribution of the burden of health care payments. The poor would have the right to demand favourable treatment in contributing to the health care financing system. This in some way overlaps with the compensation principle. The principle of fitness relates largely to issues of efficiency and the maximization of utility. This may be expressed in two conceptually different ways – sum-fitness and efficiency fitness (Moulin, 2003). Their application to the distribution of the burden of health care payments is not particularly clear. Moulin (2003) discusses these in the light of ‘resources’, his argument being that “[r]esources must go to whomever makes the best use of them” (Moulin, 2003 p.23). For example sum-fitness relates to the maximization of the sum of utilities (maximising $\sum_i U_i$) while efficiency fitness requires that resources be distributed according to how efficiently the recipient can make use of them. In some strict sense, this may require that only those who can afford to pay for health care should pay and receive health care and those who cannot afford to pay, should not receive health care.

Following this introduction, in the next section some of the better known basic theories of justices are introduced and discussed. Because this is not a philosophical analysis, the depth of coverage will only be limited to aspects that are related to economics and also to the topic of the dissertation. The theories of justice that are considered here include the utilitarian, egalitarian, libertarian, Rawlsian maximin and envy/no-envy principles. These theories are related to the discussions in this subsection.

3.4 REVIEW OF THE BASIC THEORIES OF JUSTICE

The broad notion of justice can be divided into two distinct but linked principles which were first identified by Nozick (1973) who distinguished between *historical* and *end-state or end-result* principles. This was done because of the concentration of most of the theories of justice on the current time-slices of a distribution (i.e. end-state) without focusing on the procedural aspects of justice (Nozick, 1973, Varian, 1975). The realization of this distinction has increased focus not only on how a *distribudum* is currently distributed but also on how the current distribution evolved over time and the extent to which the procedures that produce the outcome are just (Olsen, 1997). End-state principles basically provide an indication of the fairness in outcomes rather than the fairness of means and processes (Mooney, 2000, 2003). Many economists have noted the strong appeal of procedural justice (Mooney and Jan, 1997, Mooney, 2000) because of its recourse to procedures and history.

3.4.1 Utilitarianism

Utilitarianism is a moral principle of justice (Scarre, 1996) and a leading consequentialist theory of ethics, with Jeremy Bentham and John Stuart Mill as the pioneers (Konow, 2003). It generally requires a combination of four elements (Sen, 1988 p.38-39).

1. **Utility:** this is the only thing that matters and should be maximised and disutility should be minimised;
2. **Welfarism:** that the goodness of a state of affairs should be a function only of the utility information regarding that state;
3. **Sum-ranking:** that utility information regarding any state should be assessed by looking only at the sum-total of all the utilities in that state; and
4. **Consequentialism:** that every choice (of actions, institutions, motivations, rules, etc.) should be ultimately determined by the goodness of the consequent states of affairs (end-result).

Utilitarianism judges an action as right if and only if it produces more utility (welfare or well-being) overall than any alternative action (Scarre, 1996). In this regard, it is concerned with the

maximization of total utility (i.e. $\max W = \sum_i U_i$). There are several variants of this⁸. The distinction between them is often based on how utility is defined and maximised (Gandjour and Lauterbach, 2003). The idea behind sum-ranking has received criticism from theorists such as Rawls and followers of the egalitarian school of thought (Rawls, 1971, 1999). It has been argued by some (for example Rawls, 1999, Donaldson *et al.*, 2004) that utility is compared descriptively and is insensitive to inequalities in its distribution and that thereby it lacks some moral merits. Others have argued that while that criticism is correct based on the classical notion of utilitarianism (i.e. maximising the sum of individual utilities ($W = \sum_i U_i$)), it is not true for some more generalised social welfare functions (Feldstein, 1976).

Welfarism for the utilitarians requires that: where there are two states of affairs identical in individual utility characteristics, these must be judged equally good irrespective of differences in other non-utility characteristics no matter whether these other characteristics are useful in moral regards (Sen, 1979). That amounts to the notion of efficiency with no regard for equity. In fact utilitarianism has been criticised as being incompatible with equity (Stiglitz, 1982) including horizontal equity (Feldstein, 1976, Stiglitz, 1982). Even if utility can be measured, it is argued that “it does not fully capture what is relevant for equity considerations” (Roemer, 1996 p.120). Utilitarians have also been criticised for being interested only in ranking states of affairs and not addressing the problem of equitable allocation. Sen (1979 p.478) also criticises the utilitarian approach by noting that any outcome morality that incorporates the equal treatment of equals or the elimination of exploitation or priority setting requires essentially more than just utilities but the use of non-utility information as well. Other critics of utilitarianism argue that it ignores and pays no attention to individual freedom in aggregating utility or welfare. It simply assumes that when the total sum of happiness is greatest (Scarre, 1996), welfare is maximised. Some actions such as, for instance, slavery (Olsen, 1997) might increase total utility but undermine individual freedom.

⁸ Gandjour and Lauterbach (2003) summarises the various ‘types’ of utilitarianism to include act utilitarianism, rule utilitarianism, preference utilitarianism and negative utilitarianism.

In the economics of health and health care, the concepts of quality-adjusted life years (QALYs)⁹ and disability-adjusted life years (DALYs) are recognised as compatible with the traditional notion of utility (Wagstaff, 1991, Olsen, 1997). Here, the application of the greatest happiness principle of John S. Mill (Scarre, 1996) entails either maximization of QALYs or minimising DALYs.

⁹ Some economists measure the health of a single individual in terms of quality-adjusted life years (QALYs) and a community's health as the aggregation of QALYs. The QALY score of an individual is computed by weighting each remaining year of life of the individual by the expected quality of life in that year. The expected quality of life is computed as a probability-weighted average of the quality-of-life scores that are associated with each of the possible health states in which the individual might live. The scores are measured on a scale from zero (dead) to one (perfect health) and they are invariant across individuals in the same health state (Wagstaff, 1991). In this regard, QALYs are likened to the notion of utilitarianism. While there are still debates on the appropriateness of QALYs for measuring and quantifying the health of individuals and of populations (Mooney, 1989, Gerard and Mooney, 1993), their use is associated with the notion of extra-welfarism as distinct from welfarism (Culyer, 1989, Brouwer *et al.*, 2008). This is because it emphasises – in practice at least – the maximization of health. There are general disagreements as to what extra-welfarism actually means and how it should be implemented (see for example Hurley, 1988, Mooney, 1989, Culyer, 1994, Weinstein and Manning Jr, 1997, Dolan, 1999, Birch and Donaldson, 2003, Brouwer *et al.*, 2008, Coast *et al.*, 2008). For example Hurley (1988) links extra-welfarism to the maximization of solely health and not utilities; Culyer (1994) proposes an encompassing understanding of extra-welfarism that includes all 'relevant' outcomes; Birch and Donaldson (2003) discuss this in relation to Amartya Sen's capability approach; Mooney (1989) argues for components such as information and decision making to be included in utility functions of individuals and not just health; and Dolan (1999) for instance relates extra-welfarism to attaching explicit equity weights to outcomes. However different these understanding may be, Brouwer *et al.* (2008) notes that the central issues that should distinguish extra-welfarism from welfarism are the replacement of individual utilities with something else (e.g. Sen's idea of capability, broadening the definition of outcomes, etc.) and explicitly incorporating some notion of equity all of which is at variant with the traditional notion of welfarism. Some economists and philosophers have also argued, and there is reason to believe in the argument, that the use of QALYs is similar to the notion of welfarism and will lead to unjust and unfair allocations (see Wagstaff, 1991). Further, QALYs may not fully measure an individual's health. This is because they may not, in some instances, reflect the desires of the population including their right. This has also been discussed in the context of welfarism in general (Roemer, 1996). There are also controversies with the use of DALYs in measuring the burden of disease. Like QALYs, DALYs are a combination of time lived with a disability and the time lost as a result of premature mortality (Murray, 1994). It is noted that DALYs are narrow as they do not reflect differences in individuals' ability to cope with morbidity. Also, it does not include the economic cost of illness (Anand and Hanson, 1997). Several other assumptions limit the usefulness of DALYs in measuring health or the burden of disease. In summary, the principle of aggregate DALY-minimisation is inequitable in allocating resources (Anand and Hanson, 1997).

3.4.2 Egalitarianism

In general it may be argued that the basic idea behind egalitarianism is the notion that human beings are equal. At first glance the idea of equality of human beings might seem a plausible basis on which to distribute resources, or welfare. This however might lead to absurd outcomes (Olsen, 1997). This is because different human beings might require different levels of resources. This basic notion, relates particularly to traditional egalitarianism and does not take into account individual freedom or capabilities (Sen, 1985) or differences that exist among people. Egalitarianism may be general or sometimes specific. Irrespective of this, we may distinguish two forms – strong and the Rawls-type (difference principle) egalitarianism (Elster, 1992). Strong egalitarianism is a situation where everybody gets an identical share of the *distribuendum* (e.g. resources) irrespective of other considerations. The Rawls-type egalitarianism on the other hand permits inequality in the distribution of the good insofar as the distribution benefits the worst-off (the most needy) or if there is no other way to further secure benefits to the worst-off (Feldstein, 1976, Roemer, 1996, Olsen, 1997). In a broader sense egalitarian principles could involve the distribution of equal opportunities and ensure that people have equal rights and freedom (Sen, 1985, Hansson, 2004). These views are central to Sen's conception of capabilities and functionings (Sen, 1985). For Sen, the distribution of resources reflects the outcomes of choices and lack of choices that individuals may have and as such should not be a reasonable place to start. It is true that a "higher level of capability is conducive to deriving resources needed to elevate one's economic welfare and yet not all with higher capabilities choose to do so". This is because of considerations like altruism, culture, religion, or even life-style (Wagle, 2008 p.61). Some of these issues have been captured under some of the philosophical ideologies of the egalitarians.

Egalitarians broadly speaking, and unlike the utilitarians who focus mainly on the greatest happiness rule, assume that the preferred solution to the issue of fairness is that which has the most 'equal distribution' of a *distribuendum* (Olsen, 1997). To provide an understanding to how the *distribuendum* is distributed under the egalitarian philosophy, the questions posed by Roemer (1996 p.119) related to the *equalisandum*¹⁰ (Cohen, 1989) are relevant. He writes: "If one is an egalitarian, what should one want to equalize? Opportunities or outcomes? Resources or welfare?.... These positions are usually conceived to be very different." The differences in conception follow from differences in egalitarian theories. For example Rawls (see Rawls,

¹⁰ This is the object that ought to be equalised.

1971, Rawls, 1999) as noted earlier proposes the ‘difference principle’ which aims to equalise the bundle of primary goods available to each person paying particular attention to the worst-off. For Rawls, every arrangement is evaluated in terms of the interest of the least advantaged or worst-off group (Olsen, 1997)¹¹. This is the Rawls difference principle. Inequality can only be allowed provided that in doing so the utility of the worst off is increased (Rawls, 1971). However Feldstein would argue that if equality is of concern under the egalitarian principle, then why should we allow inequality no matter how little it may be, just in order to favour the worst-off (Feldstein, 1976). Arneson (1989) discusses egalitarianism as equality of opportunity and not necessarily of outcomes. Inequalities are regarded as favourable only if they result from equalised opportunities. There may be some corrective policies such as transfers to correct for inequalities that result from initial unequal opportunities. However Arneson (1990) in a later paper further notes that compensation (or transfers) to correct for inequalities may not be fair in all cases especially when the individual deliberately chooses, and is accountable for the acts and choices that cause the inequalities. This view is also held by Fleurbaey (2001) who argues for the assessment of equality in a society to be based on equality of opportunity with attention being paid to the actual content of the opportunities. Fleurbaey (2001 p.528-529) for instance writes: “I do not think that egalitarians should feel committed to the statement that equality of opportunities is sufficient for equality to be achieved. A society with equal opportunities can be replete with inequalities which no concern for personal responsibility can justify, and can totally fail to implement the egalitarian program about high-quality social relations.”

Ronald Dworkin on the other hand, in his two part thesis, after critiquing equality of welfare as an equity principle, proposes an egalitarian theory that relates mainly to equalising resources. These are defined broadly to include transferable (e.g. money) and non-transferable resources (e.g. talents, handicaps, etc.) (Dworkin, 1981a, b) and they are identified as distinct from just welfare or utility (Roemer, 1996). While it may be reasonably easy to equalise transferable resources, there are challenges with non-transferable resources. These relate to some physical and ethical constraints. However Dworkin proposes a hypothetical ‘insurance market’ as the mechanism for equalising resources. On this “Dworkin argues that a reasonable way to implement equality of resources is to calculate the postinsurance income that would have come about, had people been able to insure themselves against a bad draw in the lottery for the

¹¹ This could result in lexicographic ordering as a violation of the continuity axiom of preferences.

nontransferable resources” (Roemer, 1996 p.126). Dworkin further links the insurance mechanism with some form of tax schemes. He asks: “Can we translate that hypothetical insurance structure into a tax scheme?... Is a tax scheme constructed as a practical translation of a hypothetical insurance market, which assumes equal initial assets and equal risk, a proper response to the problem of differential talents under equality of resources?” (Dworkin, 1981b pp.324-327). A progressively structured tax scheme, based on the notion of insurance and premiums was proposed. This as Dworkin also notes could be problematic to implement (Dworkin, 1981b). There are still arguments on whether the approach proposed by Dworkin constitutes a fair division of resources or even whether equality should be a principle to pursue (Narveson, 1983a, b, Cohen, 1989) but Dworkin’s claim is that his hypothetical insurance constitutes a definition of equality of resources and not an ethical mandate (see Roemer, 1996). This discussion shows that among the egalitarian philosophers, there are still disagreements as to what exactly the distribuendum should be and this justifies the questions posed by Roemer (1996). In the context of health care financing, we may further ask if we are to equalise payments, (dis)utility from payments, post payment incomes or equalise share of income spent on health care.

3.4.3 Rawls’ maximin principle of justice

The Rawlsian maximin principle of justice and fairness (Rawls, 1971, 1999) and the egalitarian principle can be described as being related and linked (Feldstein, 1976). As noted earlier, Rawls, like some other philosophers, accepts the existence of inequality provided that the inequality is to the benefit or advantage of the worst-off in the society (compare with the discussion by Arneson, 1989 on equality of opportunity and compensation). Rawls (1974) argues that any distributive criterion should serve as a public principle that citizens, stakeholders and all concerned must understand and have some confidence in its being realised.

Rawls links the maximin equity criterion to the social contract theory. For the maximin principle of justice therefore a just outcome arises as that which will be unanimously agreed to in an appropriate initial solution that is considered fair between individuals who are conceived as free and equal (Rawls, 1974 p.141). Freedom and equality exist if individuals are deprived

of certain morally irrelevant information¹². Rawls interprets this as being behind the ‘veil of ignorance’. Rawls argues the outcome reached behind this *veil of ignorance* would involve a distribution of his so-called ‘primary goods’ that would be to the advantage of the worst-off. (These primary goods comprise liberties and opportunities, income and wealth and their distribution (see Rawls, 1974, McGuire *et al.*, 1992)). The maximin criterion, according to Rawls, is not meant to be applied to small-scale situations. It is a ‘macro’ and not a ‘micro’ principle (Rawls, 1974) such that its application is useful across major institutions of a society.

This maximin criterion has continued to receive much attention from economists. Its earliest application to taxes include those by Atkinson (1973) and Feldstein (1973) to the problem of optimal income taxation and the equity implications of progressivity of income tax (see, Rawls, 1974). Rawls (1974 p. 141) however noted that, even though the application of the maximin principle in that context might be in order, he was not in a position to examine the merits of the criterion from the standpoint of economic theory and that “whether the criterion is a reasonable distributive standard depends importantly on the sort of examination that only economists can undertake.”

Rawls (1974: p.145) also linked the maximin equity criterion to the basic foundation of communism – ‘from each according to his ability, to each according to his need’. But this is likely to infringe upon the freedom of the better endowed as they are often made to accept less than what they might have obtained under, for instance, the utilitarian principle. As argued by Rawls (1974), this can however be resolved by having societies allowing the better endowed to improve their situations but only when this is favourable to others in the society. In this regard inequalities that could be permitted should also be consistent with everyone’s self-respect.

In criticising the utilitarian view of justice and fairness, Rawls noted in his formulation that, once the least advantaged group is identified, less information is required to apply the basic maximin equity criterion than the maximization of average utility implied by the utilitarians (Rawls, 1974). However the identification of the least advantaged may prove no easy task. Some critics of Rawls, particularly those interested in health, have further argued that Rawls’ theory is based only on the notion of ‘primary social goods’. These are by no means all-

¹² The irrelevant information include their place in the society, their class, position or social status, their place in the distribution of natural assets and abilities, their deeper aims and interest, or their particular psychological makeup.

encompassing and they explicitly exclude health. However Rawls (1974: p. 142) argued that everyone has normal physical needs so that health care can be included in a broader definition of primary goods. Rawls further noted that this could still present some difficulties because maximising the least advantaged in health could impoverish the society. A related criticism is that only the worst-off within the society are considered and nothing is said about the situations of other groups within that society (see Mooney, 1994). Yet the needs of these others are equally relevant in defining and operationalizing justice and fairness. On the issue of health and health care, to make them compatible with Rawls' formulation, they may be categorised as 'basic institutions' that provide fair equality of opportunity (see McGuire *et al.*, 1992).

Whatever, it is clear that equity in health care financing implies some fairness in contribution such that some lesser burdens are placed on those who are worse-off and not just the worst-off.

3.4.4 Envy principle

Another principle of distributional fairness proposed originally by Foley (1967) relates fairness to envy and non-envy. Foley argued that there is fairness in allocation of commodities if no person is envious of the commodity bundle of another person. This position has been applied and expanded upon by some economists (see for example Feldman and Kirman, 1974, Pazner and Schmeidler, 1974, Varian, 1974). In a related formulation, Varian extended this notion of fairness and envy by comparing the allocation of goods among agents into which he brings the notion of efficiency. He writes: "[c]onsider the problem of dividing a fixed amount of goods among a fixed number of agents. If, in a given allocation, agent i prefers the bundle of agent j to his own, we will say i envies j . If there are no envious agents at allocation x , we will say x is *equitable*. If x is both Pareto efficient and equitable, we will say x is *fair*." (Varian, 1974 p.63).

Sen argues that this notion of envy and fairness may not be compatible with the absence of inequalities in, for instance, wellbeing. This is because comparisons are situational and not universal since such comparisons do not consider interpersonal variations in wellbeing functions. For example Sen (1988 p.36) proposes that "[i]f $W_1(x)$ and $W_2(x)$ are the well-being functions respectively of persons 1 and 2, and $W_1(x_2) > W_1(x_1) > W_2(x_2) > W_2(x_1)$ when x_1 and x_2 are the bundles respectively enjoyed by persons 1 and 2, then it is person 1 who envies person's 2's bundle while person 2 has no such envy, even though it is person 2 who is clearly worse off in terms of well-being than person 1. Further, if we have $W_1(x_1) > W_1(x_2) > W_2(x_2) >$

$W_2(x_1)$, then the condition of ‘nonenvy’ is completely satisfied on both sides, even though person 2 is clearly worse off than person 1. Since economic misfortunes are often associated with variations in the well-being functions, due to such factors as disability, illness, age, social discrimination, etc., there is much scope for doubting the fairness of the fairness criterion.”

Some critics also argue that the principle of envy in assessing fairness reduces the issue of distribution to the consideration of individuals (Mooney, 1996). Varian has however provided an extension to the envy principle to go beyond such comparisons of individuals. He proposes the concept of *coalition fairness* to incorporate groups of agents. Here an allocation is defined as coalition fair if and only if no coalition of a given size prefers any aggregate bundle of any coalition of the same size or smaller (Varian, 1974 pp.75-76). However the operationalization of this as a principle of distributive justice still poses some difficulty. This arises because of issues surrounding how to measure coalitions and their respective sizes. Moreover we should be ideally concerned about communities and a group of people rather than the formation of coalitions. Also others have discussed the ideas of altruism and solidarity disposition (Crocker, 1977) that are of importance in how societies respond to the inequalities of others. Such ideas are at least incorporated in the broader idea of envy and also the extension to the coalition fairness.

3.4.5 Libertarian principle of justice

The libertarian principle is generally regarded as some sort of political theory or philosophy that is based on the principle of neutrality (i.e. there should be a neutral relationship between the state (and its laws) and the various conceptions of what is good as held by individuals) (Alexander and Schwarzschild, 1987). According to Alexander and Schwarzschild (1987 p.86), libertarians can be distinguished by understanding the resources that are up for ‘neutral’ distribution and what the proper formula for distributing those resources is. Libertarians maintain that the right to freedom and self-ownership is important in defining what is equitable. These include absolute respect for the right to private property that recognises both the positive and negative freedoms¹³ that individuals have (Denier, 2007). One of the widely noted libertarian theories (that is explained in detail later) is Robert Nozick’s entitlement theory

¹³ Negative freedom is basically freedom from interference. This means that there should be no hindrance in the way of any individual in achieving any goal. Positive freedom relates to “the claim to something, the possibility to be someone, to pursue and realise a private goal in a certain manner, the possibility to be aware of choices and to be able to explain them in relation to ideas and goals” (Denier, 2007 p.45).

(Nozick, 1973, 1974). The starting point of that theory is John Locke's ideology (i.e. the Lockean proviso) of self-ownership such that "[b]y virtue of self-ownership, a person is entitled ... to appropriate objects in the external world as his own, as long as he leaves 'enough and as good in common' for others" (see Moulin and Roemer, 1989 p.348).

For the libertarians, even if it can be argued that health care would contribute to individual freedom, mandatory taxation or public health insurance contributions to organise a progressive health care financing system would be viewed as a libertarian injustice (Denier, 2007). This is because it constitutes partial ownership of others' rights and actions (Nozick, 1974). Here only private health system and insurance that is based on voluntary redistribution can be accepted. Because the market mechanism is favoured over any other form of intervention, health care is seen as being like any other commodity in the market. Therefore the distribution of health care and the burden of health care payments should be left to market forces.

In general the libertarian philosophy is a 'hands-off' theory that places emphasis on the rights and freedoms of individuals and the avoidance of violation of these rights (Denier, 2007). Some critics note that community or society values rather than individual values should be emphasised as different societies have their own meaning of the good to be distributed and should determine their proper distribution (Lamont, 2004). This view is commonly held by the communitarians. Even within the libertarian ideology, it is important to note that having rights does not automatically translate to the availability of, for instance, social services or goods that improve wellbeing (Denier, 2007). There is need for some sort of intervention to address inequalities. While some libertarians accept this to some extent, most stick to the *hands-off* philosophy.

3.4.6 Nozick's Entitlement Theory

Nozick (1973, 1974) proposes that justice should be based on 'entitlements' or 'holdings'. Based on the Lockean proviso he began his presentation by distinguishing between *historical* and *end-state or end-result* principles. Nozick argues that most theories of justice (i.e. distributive justice) are ahistorical. The historical aspects deal with how the current distribution has evolved (Varian, 1975). The concept of justice in holding (or entitlement) according to

Nozick consists of three parts: *original acquisition of holdings*¹⁴, *transfer of holdings* from one person to another¹⁵, and *the rectification of injustice in holdings* that seeks to correct past injustices that now cause the present injustice (Nozick, 1973). Nozick writes that “[t]he general outlines of the theory of justice in holdings are that the holdings of a person are just if he is entitled to them by the principles of justice in acquisition and transfer, or by the principle of rectification of injustice (as specified by the first two principles)¹⁶. If each person’s holdings are just then the total set (distribution) of holdings is just” (Nozick, 1973 p.49). For Nozick injustice or inequity depends on how it came about and should be traced to history. A distribution can be described as just only if it arises from another just distribution through legitimate means (Nozick, 1973). However Nozick was hesitant to propose a theory of justice based on entitlement. He writes: “[t]o turn these general outlines into a specific theory we would have to specify the details of each of the three principles of justice in holdings ... I shall not attempt that task here” (Nozick, 1973 p.49).

Nozick’s conceptions have been criticised by Varian for example who argues that “how an allocation was reached has *something* to do with justice; but it is just not the whole story. ... If a process can be radically affected by turns of fortune that are arbitrary from a moral point of view, it seems unreasonable to attach great moral significance to the outcome of such a process” (Varian, 1975 pp.286-287). Further, as opposed to the end-state theories that “decide what a perfectly just state is and then try to move toward it”, Nozick’s theory pays little attention to rectifying the current distribution (Varian, 1975). It is described as anti-redistributive because “[c]oerced redistribution would be an unjust redistribution of private property, by illegitimately considering it as public property” (Denier, 2007 p.47). Redistribution can only be permissible, voluntarily through acts of charity which should be viewed differently from any principle of justice. Even with the principle of rectification, it needs still to answer the question:

¹⁴ This is “the appropriation of unheld things. This includes the issues of how unheld things may come to be held, the process(es) by which unheld things may come to be held, the things that may come to be held by these processes, the extent of what comes to be held by a particular process, and so on” (Nozick, 1973 pp.46-47).

¹⁵ This is related to Moulin’s conception of the principle of transfer discussed earlier.

¹⁶ i.e. the principles of holding and transfer

what would the outcome be if the injustice had not occurred? This requires adequate specification of the initial endowments. How then do we trace back to this initial position?¹⁷

Currently, we live in a world described as unfair with inequalities in health and access to health care (see for example Culyer and Wagstaff, 1993, Mooney, 1996, Deaton and Paxson, 1998, Coovadia *et al.*, 2009); inequality in income (see for example Shorrocks, 1978, Deaton and Paxson, 1998); inequalities in access to social services and community resources (see for example Blackburn and Jarman, 1993, Logan and Bian, 1993, Klasen, 1997); and unfair distribution of health care financing burden, particularly paying for health care out-of-pocket (see for example Castano *et al.*, 2002, McIntyre *et al.*, 2006). While some of these are based on current distributions, historical contexts have been important in understanding some of these distributions. This can be seen in the context of countries such as Australia with the aboriginal Australians (Mooney, 1996), and South Africa with the apartheid system (see for example Klasen, 1997).

In conclusion, the quote from Musgrave captures, in a nutshell, the intricate nature of using different theories of distributive justice to underpin the assessment of progressivity and distributional impact of taxes and health care payments. Musgrave (1996 p. 348) writes:

“Economists ... can be useful in setting forth the analytics that help determine the “correct” pattern of progression or vertical-burden distribution under alternative views of distributive justice, but they cannot pronounce on which is *the* correct one.”

Musgrave goes on to highlight the value-laden nature of the theories of distributive justice and notes some of the limitations that economists may face in applying them. He writes that the inability of economists to specify the ‘correct’ distribution that can be deemed fair is a reflection not only of

“our limitations in adequately measuring economic responses to taxation but also, and more importantly, that the outcome depends on the underlying equity norm, a choice which is not to be captured within the confines of Paretian efficiency. This choice depends on how the problem of distributive justice is

¹⁷ Varian (1975 p.236) writes extensively on this issue. For instance he asks: “[a]re we supposed to trace back ownership of American land to the Indians and then try to rectify the wrongful appropriation of it by the methods Nozick discusses?” This is a very difficult task.

viewed – for example, from a Lockean entitlement perspective, from rules of equality in sacrifice, or from the utilitarian target of maximum welfare (least total sacrifice) with its ethical anchor in the premise of impartiality. All this [*sic*], to be sure, may be taken as subsumed by the stipulated shape of the social welfare function, and the young economist, eager to get on with maximizing, may be impatient with the underpinnings. But to understand what goes into the underlying reasoning, and hence the case for or against progression, the various inputs need to be sorted out.”

Such inputs as noted by Musgrave may be the values of the community. Musgrave accepts that public perception on tax and payment equity may not coincide with those outlined by philosophers and economists (Musgrave, 1996). The writing of economists like Mooney (1996 p.102) is relevant here. He writes that:

“Here perhaps lies the solution to how to answer the question of ‘equality of what?’ with respect to both horizontal and vertical equity. *Let us ask the community*” [emphasis added].

Mooney explains further that the question of equity (especially vertical equity) can be addressed mainly through establishing community preferences on what weights be attached to different groups and outlining the basis upon which such discrimination should occur or not occur. Mooney reports on a preliminary survey on “preferences among public sector decision makers (presumably on behalf of the community)” in the context of health care and this “suggests that age groupings, socio-economic groupings, groups with different existing health status (but not gender) might form appropriate bases for such discrimination.” While there may be differences across communities, Mooney advocates for the extension of such surveys to cover the community directly.

Based on the outlined theories, different approaches can be delineated that underpin the assessment of equity in health care financing. The next section discusses these approaches bearing in mind the various theories of distributive justice and how they relate to the approaches. These theories also have implications for how redistribution can be conceived.

3.5 APPROACHES TO ASSESSING EQUITY IN HEALTH CARE FINANCING

“The prominence of the tax progressivity issue puts economists in an uncomfortable position, because we know in our heart that progressivity is not solely an economic question. It is equally a matter of ethics, or moral philosophy, because it involves choosing between situations where some people are better off and others worse off. Economics cannot settle which people are more deserving. Economic reasoning cannot determine whether it is good idea to take one dollar from a wealthy family in order to give one dollar – or, even more problematically, fifty cents – to a poor family” (Slemrod, 1996 p.1).

Even though the above quotation was discussed in the context of general taxes, it is also relevant to the assessment of equity in health care financing. However difficult this subject matter may be, in the economic literature three approaches to assessing equity in health care payments may be distinguished. These have been identified and discussed in the context of tax equity (Musgrave, 1996, Slemrod, 1996). These include (i) the benefit principle (or *quid pro quo*), (ii) the sacrifice rule, and (iii) the ability-to-pay principle. These will now be examined in turn. They are based on different notions of fairness introduced in section 3.4.

The benefit principle originated from the defence sector and was then later applied to other issues such as taxes in public finance. It is concerned with the concept of fairness in exchange or *quid pro quo* (Musgrave, 1996). The principle suggests that individuals should be taxed (or make payments) based on the benefits they receive from government expenditures (Hines Jr, 2000). For example in the absence of government protection services, such as the police and military, fire-services, etc., holding properties would be riskier. Because these services serve to reduce risks, those who own large amounts of properties would be required to pay more than those who own fewer or no properties (Miller, 2005). Under this approach, the distribution of benefits received is traditionally based on Lindahl prices or tax (i.e. the situation where marginal benefits received by an individual determines the price paid) or on linear pricing of public goods (Hines Jr, 2000). While Lindahl prices may vary across individuals and could lead to an inconsistent and paradoxical computation of benefits, the linear pricing “ensures that consumers unanimously prefer the efficient level of public good provision, and that all consumers are made better off by reduced costs of providing public goods” (Hines Jr, 2000

p.484). This pricing mechanism is built on the traditional benefit incidence analysis (BIA) that measures the monetary benefits (or subsidies) individuals get from the use of goods and services provided by the government. Benefit taxes are then set to cover the cost of providing these goods. Thus the benefit principle is based, to a large extent, on some efficiency principles and does not consider differences in capacity or ability of individuals to contribute to the system.

In this context it is important to note that some goods provided by the government are such that the poor should use more and benefit more. If this is the case, the poor would be required to contribute more than the rich. Even when the use of some goods provided by the government is pro-rich, when payments are structured in this way, while the rich may pay more in absolute terms they may pay less proportionally (i.e. according to their ability to pay). It is then not surprising therefore that equity in health care financing generally does not adopt the benefit principle. However, private health care financing may do so.

The second principle, the sacrifice rule, is rooted in ‘sacrifice theory’. It is essentially egalitarian in orientation. Here taxes (including other compulsory health care payments) are losses or sacrifices that individuals and households make (Musgrave, 1996, Miller, 2005). According to this rule decision makers are then required to distribute the sacrifice ‘equitably’ among the population. Thus the law of equality of sacrifice requires that all should be treated equally (Musgrave, 1996) but then equality here can be interpreted in many ways: equality of absolute sacrifice, equality of proportional sacrifice, or equality of marginal sacrifice (Musgrave, 1996). Equality of absolute sacrifice requires that everyone gives up the same amount of utility (i.e. for any given income X and health care payment or tax $t(X)$, $U(X - t(X)) - U(X) = k$ for all X). Equal proportional sacrifice requires that each individual should give up equal share of total utility derived from money (i.e. $U(X - t(X)) / U(X) = k$ for all X). Equality of marginal sacrifice requires that on the margin, the additional utility given up by each person is equal (i.e. $\partial U(X) / \partial t(X) = k$ for all X) (Miller, 2005). Because of diminishing marginal utility of income¹⁸, progressive tax rates could result in both proportional or progressive reductions in utility (Miller, 2005). Proportional reductions in utility will arise when the progressive tax structure is set such that richer tax payers give up the same level of utility as poorer tax payers.

¹⁸ Here declining marginal utility of income is assumed (i.e. for a given amount of money, the value to the poor is greater than for that for a rich person. This also implies that the utility of money declines as income increases (Miller, 2005)).

Progressive reductions arise when the reductions in utility associated with the progressive tax structure is less among poorer tax payers than richer tax payers.

Equal absolute sacrifice and equal proportional sacrifice were the earliest forms of the sacrifice rule but due to the appeal of the equal marginal sacrifice and the rise of the marginalist school of thought, this last became more widely accepted through the works of Francis Edgeworth and Arthur C. Pigou (Musgrave, 1996). (Equal marginal sacrifice is also termed least-total-sacrifice (or maximum welfare) (Musgrave, 1996).) In effect it is associated less with fairness than with efficiency. This “move from Mill’s equal absolute or proportional sacrifice to Pigou’s enshrinement of equal marginal sacrifice thus involved a paradigm shift from equity to Pareto efficiency as the basic criterion” for the distribution of tax or payment burden. The marginalist approach that uses the diminishing marginal utility is appealing because “taking a dollar from a person with the larger income involves less sacrifice than taking a dollar from a person with the smaller income. If required, a second dollar might still entail less sacrifice from the person with the higher income than the person with lower income if after the second dollar the higher income person were still richer than the other person” (Miller, 2005 p.441). The efficiency twist here, according to Pigou, requires taking from the richest person’s income “in the form of maximum progression and absorption of top incomes until the required revenue is obtained” (Musgrave, 1996 p.345) and, where necessary, incomes are provided out of taxes to the poorest.

One of the difficulties with the sacrifice approach is the quantification of utilities and the specification of utility functions. Utility (including utility of money) does not depend on observable factors alone but also unobservable ones. In fact utility itself is unobservable. Relating health care financing to unobservable utility functions becomes problematic and it is this which reduces the practical relevance of the sacrifice rule in the field of equity in health care financing. In empirical applications, the approach has to rely on postulations about the shape of the social welfare function and empirical propositions regarding the responses of payers (Musgrave, 1996). Even if these did reflect practical realities, the approach would still prove difficult to implement.

While the benefit principle is tied to some form of reciprocity from the government and the sacrifice principle is egalitarian linked to equalising utilities in one form or another, both, to a considerable extent, draw on efficiency criteria. The ability-to-pay principle is based explicitly

on structuring payments and contributions in relation to an individual's ability to pay. There are debates as to what the appropriate measure of ability-to-pay should be (Buehler, 1945, Wagstaff and van Doorslaer, 1993, Miller, 2005). For instance J.A. Hobson defines it as "the extent of the person's ability to create a surplus above the cost of producing income" (see Miller, 2005). Wagstaff and van Doorslaer (1993) note that ability to pay could be measured by pre-tax (i.e. gross) income or pre-tax income plus imputed income from physical assets such as the individual's or family's house. Ability to pay has also been measured by total expenditure, consumption, or pre-tax income (or consumption) less expenditure on basic necessities (e.g. food, clothing, shelter, housing, etc.) (World Health Organization, 2000). In empirical applications, however, most studies use pre-tax income or consumption expenditure due to difficulties in quantifying, for instance, the value of household assets (Wagstaff and van Doorslaer, 1993). Whatever, the ability-to-pay principle requires that those with greater ability to pay, contribute more to the health system than those with lesser ability to pay (Wagstaff and van Doorslaer, 2000). It is also to be noted that such payments should not be linked to the utilization of health services (Wagstaff and van Doorslaer, 2001).

This approach brings the fair treatment of those with a lesser ability to pay into explicit focus. It is this approach that is adopted in this thesis. It is important to recognise that simply to note that those with a higher ability to pay (ATP) should pay more could result in a progressive, proportional or regressive payment structure when the ratio of payments to ATP are, on average, an increasing function of ATP, constant function of ATP or decreasing function of ATP respectively (Wagstaff and van Doorslaer, 1993). It therefore becomes important to qualify the ATP principle as a progressive principle (i.e. payments as a proportion of income is an increasing function of ATP). Regressive forms of payments are normally regarded as inequitable (Wagstaff, 2002b) and unfair.

Generally in the health care financing literature, as also adopted in this thesis, the notion of equity in health care financing is linked to financing health care in a progressive fashion according to ability to pay (Culyer and Wagstaff, 1993, Wagstaff, 2002b, Donaldson *et al.*, 2004, Cissé *et al.*, 2007). On this economists like Wagstaff (2002b p.114) reacting to the World Health Organisation's proposed fairness in financial contribution index notes that "most policymakers feel comfortable with the ability-to-pay principle as the underlying principle of health care finance" and it is "most unlikely that most – if any – interpret this in terms of a hard-and-fast rule on proportionality." However, it is rarely the case to have policy-makers

specify either how ATP is to be defined or the preferred progressivity relationship between payments and ATP (Wagstaff and van Doorslaer, 1993, Wagstaff and van Doorslaer, 2001).

Broadly speaking, the recent literature on equity in health care financing is based on two related principles – the horizontal equity (*HE*) principle and the vertical equity (*VE*) principle. These are closely linked to the ability-to-pay principle. The *VE* principle involves the unequal but equitable treatment of unequals or reducing welfare gaps between unequal individuals (Culyer and Wagstaff, 1993, Mooney, 1996, Duclos and Lambert, 2000, Duclos *et al.*, 2003). In health-care financing this is linked to the analysis of progressivity and it requires that individuals, households or groups with different abilities to pay (or income) pay appropriately different amounts for health care. The *HE* principle requires that individuals or groups with the same ability to pay (or who share the same level of welfare before any health-care payments) make the same contribution to health care (Culyer and Wagstaff, 1993, Wagstaff and van Doorslaer, 2000, Cissé *et al.*, 2007). The operation of the *HE* principle is rarely questioned because of its intuitive appeal. However policy analysts disagree on the degree to which the *VE* principle could and should be implemented and how great any differential payments would be that would be described as vertical equity (Mooney, 1996 p.101). Because this is largely an issue of normative judgement, decisions on this are likely to vary from society to society depending, for example, on how compassionate a society is. It can be argued that the *HE* principle hinges on the moral idea that all human beings are worth the same. This is what drives the equal sacrifice theory of taxation (Duclos and Araar, 2006). Another notion of equity in taxation and also health care financing is the absence of re-ranking of individuals or households as a result of health care payments (Wagstaff and van Doorslaer, 1997, Gerdtham and Sundberg, 1998, Wagstaff *et al.*, 1999a, Wagstaff, 2002b, Duclos *et al.*, 2003). This has also been referred to as “improper treatment of unequals” (Abu-Zaineh *et al.*, 2009). The general idea here is that households should maintain their pre-payment ranks (or positions) even after making health care payments. A violation of this leads to re-ranking of households. For example if household A is richer than household B before paying for health care, it should not be the case that household B becomes richer as a result of the payment. Similarly, if household C is as rich as household D, health care payments should not make any one richer/poorer than the other. This though related to the horizontal equity principle, could still

occur in the presence of horizontal equity (Aronson *et al.*, 1994) especially when the marginal tax or payment rate is greater than one¹⁹.

3.6 REDISTRIBUTION AND HEALTH CARE FINANCING

Up until now we have looked at, and have tried to understand, the various perspectives that may underpin how the burden associated with health care financing should be distributed across individuals, households and groups of people. This is analogous to examining how the tax burden, income or wealth is distributed. The study of such distributions is important at a number of levels but here particularly in the context of their usefulness primarily to government in formulating redistributive policies (Boadway and Keen, 2000) such redistribution being most simply understood as “an unrequited transfer of resources from one person to another” (Boadway and Keen, 2000 p.679). Because equity in health care financing relates to the ability-to-pay principle, redistribution involves taking away some resources from the wealthy and distributing these to the less wealthy. Such redistribution may extend beyond income, money or health care payments to the distribution of other social goods and services which enhance welfare (O'Donnell *et al.*, 2008b). The general rationale for redistribution has been divided in three (Boadway and Keen, 2000): the pursuit of social justice and equity; achieving mutually advantageous efficiency gains; and the exercise of the redistributor's self-interest (e.g. when the state uses its coercive power or when self-interested individuals redistribute toward themselves).

While the last two rationales may be important in their own right, of primary relevance to this review and this thesis is that related to the pursuit of social welfare and equity. This may take different forms depending on the underpinning theory of distributive justice. For instance the utilitarians, interested in the maximization of welfare ($\max W = \sum_i U_i$) based on their welfarism orientation, will attach distributional weights to individuals and use these in marginal redistributive reforms. For example if individual i has a lump sum income of X_i and corresponding utility U_i , the distributional weight for person i is given as $\beta_i = \partial U_i / \partial X_i$. Analogously the distributional weight for person j is computed as β_j . The weights imply that

¹⁹ The idea behind this is discussed under the methodological debates. It is however important to note that marginal tax or payment rates greater than 100% could distort efficiency in a system and could be counter to incentive preservation (Lerman and Yitzhaki, 1995).

the higher the value of β_i , the greater is the social gain from any increase in person i 's income (Boadway and Keen, 2000). If therefore $\beta_i > \beta_j$ (for $i \neq j$) then redistributing \$1 from person j to person i would increase social welfare. For taxes, the marginal costs of raising extra revenue may be used. For instance if for simplicity c_i is the marginal cost of raising an extra unit of revenue via tax from good i and c_k is that for good k and $c_i > c_k$, if social welfare is to be raised, given constant revenue, the tax on good i would be reduced while that on good k would be increased until marginal costs are equal (Ahmad and Stern, 1984). This is the usual first-order condition for optimization. Another approach in respect of taxes and payments is the inverse optimum approach that uses the "first-order conditions to solve for social welfare weights on increments in income to each household assuming existing taxes are optimum" (Ahmad and Stern, 1984 p.260). While this welfarist approach is based on efficiency, it does not take into account the fact that some goods are consumed more by the poor. It is important to note that raising taxes or payment rates among the poor, relative to the rich, in order to satisfy the optimum criterion may not be a populist view if there is strong concern for the welfare of the poor.

In contrast to the libertarian view (see for example Nozick, 1974), this welfarist approach ignores private property. In fact Nozick's formulation only supports redistribution if this is a voluntary act of charity. For the libertarians, any redistribution that proceeds without regard to the private ownership of rights amounts to injustice. Basing (re)distribution on income and/or utility is also narrow and problematic (Sen, 1985, 1988). For instance even after redistribution there may be political repression (Boadway and Keen, 2000) and misery (Sen, 1988). This is because things other than just utilities and income can be important such as whether people deal honestly with each other (Boadway and Keen, 2000).

Health care financing systems in general may well have some income redistributive effects (O'Donnell *et al.*, 2008b). These may be intended or unintended, pro-poor or pro-rich. Pro-poor redistribution implies that the distribution of income net of taxes and payments is 'more equal' than the distribution of gross income (Boadway and Keen, 2000). When incomes are redistributed to the rich you have a pro-rich redistribution pattern. Redistribution could also be classified as either *vertical* or *horizontal*. Horizontal redistribution occurs when individuals or households with the same level of gross income make the same tax or health care payments while vertical redistribution occurs when such taxes or health care payments are

disproportionately related to ability to pay (Lambert and Ramos, 1997, O'Donnell *et al.*, 2008b).

In empirical analyses of the redistributive effect of health care financing, this question of pro-poor or pro-rich is measured as the difference between the Gini index pre-payment income (G_X) and post-payment income (G_{X-T}) (i.e. $RE \equiv G_X - G_{X-T}$) (Reynolds and Smolensky, 1977). This approach was also used in this thesis. The redistributive effect has been decomposed into comprise three main elements (i.e. horizontal equity, vertical equity²⁰ and re-ranking) (Aronson *et al.*, 1994, Wagstaff and van Doorslaer, 1997). This decomposition is attributed to the work of Aronson *et al.* (1994) and is used in other recent analytical frameworks (see for example van de Ven *et al.*, 2001, Duclos *et al.*, 2003, Urban and Lambert, 2008). While the analysis of progressivity and the distribution of the burden of payments discussed earlier may reveal and provide insights into some aspects of vertical equity, it cannot show distinctively the extent of unequal treatment of equals (i.e. horizontal inequity) and the reranking or rank reversal that can be induced by health care payments. This is because the methods only capture how such payments vary with some measure of ability-to-pay.

It is worth noting that horizontal inequities could arise in a vertically equitable tax or health care payment system because of, for instance, variations in tax or payment rates across regions and provinces, the existence of tax deductibles for certain *categories* of individuals, differences in the incidence of illness faced, risk-rated contributions to health insurance schemes, different sources of income (e.g. formal employment vs. informal employment), etc. (Wagstaff and van Doorslaer, 1997, van Doorslaer *et al.*, 1999, Wagstaff and van Doorslaer, 2001). In the case of personal income tax, where most of these analytical techniques have been developed and applied, horizontal inequity can still occur but this is less of a concern in this context as tax rates are usually structured progressively with income (Wagstaff and van Doorslaer, 1997). In the context of indirect taxes, private insurance contributions and other payments, because these are related to the individual's tastes, choices and levels of consumption, such as sales taxes, excise taxes and value-added tax, differential treatments can arise (van Doorslaer *et al.*, 1999) and may not be closely linked to ability to pay (Wagstaff and van Doorslaer, 1997). Because there is some degree of choice involved when one considers non-compulsory payments, it may be difficult to label some of these differential treatments in this instance as 'inequitable.'

²⁰ In the strict Aronson *et al.* (1994) formulation, the vertical equity component is affected by the average payment rate defined as the average share of income devoted to health care.

However it can be agreed that an overall positive or pro-poor redistributive effect implies that health care payments induce a reduction in income inequality.

Generally, the primary objective of any health care financing system is not to redistribute income, but to generate resources for the health system (Bilger, 2008, Cavagnero and Bilger, 2010) and to ensure that households are fairly treated when financing health care. However, for health care financing to have a redistributive effect, voluntary payments and payments made at the point of consumption are to be avoided. This is most evident in general taxation and in some forms of mandatorily financed health systems (O'Donnell *et al.*, 2008a, O'Donnell *et al.*, 2008b) such as the United Kingdom and France. It is reasonable in most cases, Wagstaff (2008) has argued, to treat health spending as not contributing directly to household welfare²¹ and that they be non-discretionary²².

Because redistribution in the context of health care financing is based on the assumption that payments are compulsory and are independent of health care utilization, an important question that has often been raised is whether direct out-of-pocket (OOP) payments or other forms that do not meet those criteria could generate any redistributive effect. Though OOP payments are by definition 'personal' private payments, the occurrence of illness that may warrant OOP spending is largely stochastic. This means that to some extent at least households' choices are limited with respect to OOP payments (Wagstaff and van Doorslaer, 1997). Such payments

²¹ Deaton and Zaidi (2002) also note that "such expenditure reflects regrettable necessity that does nothing to increase welfare." In general illness and the resulting payments for health care result in reductions in welfare (Deaton and Zaidi, 2002).

²² In many instances, because the occurrence of an adverse health event may be largely stochastic (or referred to as "shocks"), and there is a strong desire to get well, it is only natural that payments for such conditions can be seen as non-discretionary. Of course not all payments for health care are non-discretionary, most obviously for some types of cosmetic surgeries. Again some individuals may choose not to pay for health care perhaps because they cannot afford it. In general, however, it is not uncommon to find that households borrow to finance health care (see for example Narayan *et al.*, 2000, Knaul *et al.*, 2006) or have their bills paid by members of their immediate or extended family. Empirical investigation has also shown that income-elasticities of health care spending are generally low for developing countries (Deaton and Zaidi, 2002). Though Deaton and Zaidi (2002) estimated a mildly high elasticity ($\eta = 1.41$) for South Africa, this was based on a dated dataset which is reflective of the 'old' South Africa. Using the 2005/2006 Income and Expenditure Survey, the estimated income elasticity of health care payments was 0.40. This is even smaller than those estimated for Vietnam (0.86), Brazil (0.85), Nepal (0.75), and Panama (0.80) (Deaton and Zaidi, 2002).

could also be made to allow consumption of health care for others within a household or extended household. In a caring or altruistic system, individuals may help to finance the health care costs of others (Boadway and Keen, 2000). On this O'Donnell *et al.* (2008b p. 197) argue that while tax payments constitute some legal obligation, out-of-pocket payments can come with moral obligations. They write that “the moral compulsion to purchase vital health care for a relative is no less strong than the legal compulsion to pay taxes.” However, unlike compulsory taxes or payments, such private payments have limited (even if not zero) scope for redistribution.

In conclusion, based on the review contained in this chapter, this thesis applies the ability-to-pay principle to the analysis of the distributional impact of health care financing in South Africa. This will comprise the examination of both the horizontal and vertical equity dimensions. The major methodologies that are often used in such analysis are introduced in the next chapter.

Chapter 4

METHODOLOGICAL AND EMPIRICAL REVIEW OF LITERATURE

4.1 INTRODUCTION

This chapter reviews the methodologies and more generally the empirical literature on the distributional impact of health care financing. The first part of the chapter provides a thorough overview of methodological debates and developments in the assessment of equity in health care financing. These methodological developments follow from our understanding of equity in health care financing based on different theoretical underpinning. They also inform the nature of analyses that is performed in this study. The second part of the chapter provides detailed summary of empirical evidence from around the world. These are later compared with those from this study.

In the review, particular attention is given to the methodologies used in and the empirical analyses of the progressivity and redistribution associated with health care financing. In assessing such progressivity, it is important to bear in mind that there is no generally agreed ‘extent’ of progressivity that any health care financing mechanism or system should achieve to be deemed ‘equitable’. Regressive mechanisms are however regarded as inequitable. Similarly there is no agreement on what a preferred redistribution would look like that any health care financing mechanism or system should induce. Certainly a pro-poor redistribution is favoured over a pro-rich one.

Specifically this chapter provides a review of related methodological developments and their application to the assessment of equity but in particular looking at the distributional impact of health care financing. It also allows a critical assessment of studies conducted using these methodologies and at the same time attempts to use this to provide a guide to the rest of the study. The first part provides a summary of the methods used in assessing progressivity in health care financing while the second reviews studies that have applied these methods. The third part sets out some methodological debates around the assessment of the redistributive

impact of health care financing while the fourth reviews the literature on assessing redistribution in health care systems. While most studies in developing countries that investigate progressivity in health care financing have focused on out-of-pocket payments, most of those on redistribution have been conducted in developed countries.

4.2 MEASURING PROGRESSIVITY IN HEALTH CARE FINANCING

Assessing progressivity in health care financing involves relating health care payments to households' income or some other measure of ATP, which was earlier introduced in principle (Lambert, 2001). Again as discussed earlier, depending on what proportion of income is paid by rich and poor, the relationship can be progressive (rich higher), proportional (rich and poor the same) or regressive (poor higher). Beyond that certain key questions arise: what is the preferred relationship between income and health care payments? Should it be progressive? If so, how progressive? Such progressivity has no agreed degree (Wagstaff, 2002b), and what is 'equitable' is likely to vary from one society to another.

Using a simple framework borrowed from the analysis of taxes in the public economics literature, we denote $T_i(x)$ as the amount an individual with gross income x contributes to the health system via the health care mechanism i (e.g. direct taxes). The financing mechanism i is said to be progressive (regressive) if the elasticity of T with respect to x is greater (less) than one. If proportional, the elasticity is unity (Kakwani, 1977). Stated differently, a progressive (regressive) system is one where the average payment *rate* increases (decreases) with income. Thus except in a proportional system, the relative purchasing power of different income units are affected by health care payments (Lambert, 2001).

Mathematically, $T_i(x)$ is strictly progressive *iff* $T_i(x)/x$ increases with x for each i . Further, if $T_i(x)$ is everywhere differentiable, strict progression will imply that:

$$\frac{d[T_i(x)/x]}{dx} > 0 \quad \forall x > 0 \quad (4.1)$$

and weak progression²³ that:

²³ This criterion for weak progression takes into account the case where health care payments may for instance show a proportional relationship up until a threshold say $x > x_o$ thereafter showing a strictly progressive structure.

$$\frac{d[T_i(x)/x]}{dx} \geq 0 \quad \forall x > 0 \quad (4.2)$$

There are various methods in the literature for assessing progressivity²⁴ (see Lambert, 2001). Irrespective of the choice of method, when our interest is in assessing progressivity in an entire system, two major steps are involved. First, the relative progressivity of each financing mechanism is assessed. Secondly, progressivity in the overall system is evaluated (O'Donnell *et al.*, 2008b) taking into account the relative contributions of each of the financing mechanisms in the overall financing system.

The easiest and crudest method used by some researchers is to compare different quantiles' shares of income spent on health care and attempt to deduce progressivity (see for example Lyon and Schwab, 1995). This proceeds by categorising households into quantiles (e.g. deciles, quintiles, quartiles, terciles, etc.). For each the share of the quantile's income spent on health care via each mechanisms (for instance general tax, out-of-pocket payments, social insurance contribution, etc.) is computed. Progressivity is then assessed by considering how the ratios (i.e. the analogue of average tax rate) vary across quantiles. This however has some limitations – it is not sensitive to variations that may occur close to the cut-off points for categorising households; it does not show a holistic picture of how payments to income ratio vary across the entire distribution of income; and the exact extent of progressivity cannot be obtained simply by looking at these ratios. In fact in some situations it could be difficult to distinguish which mechanism is more progressive or regressive than the other (Musgrave and Thin, 1948, Kakwani, 1977, Suits, 1977).

In response to these limitations, some more formal indices and curves have been developed and used in the assessment of progressivity. These include some forms of the Lorenz curve (Lorenz, 1905), Gini index, and concentration curves and indices. One of the earliest is the Musgrave and Thin index (Musgrave and Thin, 1948). Other popular indices are the Suits index of progressivity (Suits, 1977) and the Kakwani index (Kakwani, 1977). The Musgrave and Thin index simply compares inequality in pre-payment income with that of post-payment

²⁴ There is often a distinction made between structural progression and effective progression (Khetan and Poddar, 1976). The former describes the properties of a tax or payment system independent of the level and distribution of incomes upon which taxes are levied. Effective progression which is of particular interest in this thesis has become popular among economists “measures the extent to which a given tax structure results in a shift in the distribution of income towards equality” (Musgrave and Thin, 1948 p.510).

income. In fact what it shows is a measure of the redistributive effect of health care financing (Kakwani, 1977). It is not a pure measure of progressivity. If inequality in post-payment income is less than in pre-payment income, under Musgrave and Thin's formulation we conclude that the health care financing mechanism is progressive. A regressive mechanism increases inequality in the post-payment income²⁵. For empirical purposes the Musgrave and Thin index is computed as:

$$\pi_{MT} = (1 - G_{X-T}) / (1 - G_X) \quad (4.3)$$

where G_X (G_{X-T}) is the Gini index of pre- (post-) payment income. Specifically when $\pi_{MT} > 1$ payments are progressive. One of the major criticisms of this is that it does not assess progressivity *per se* but redistribution and does not take elasticities of payment into account (Kakwani, 1977). On this Kakwani (1977 p.74) notes that “[i]f we assume that the change in the average tax rate does not affect the before-tax income distribution, the Musgrave-Thin measure may show an increase or decrease in progressivity even if the tax elasticity has remained constant at all income levels”.

Under Musgrave and Thin's formulation a tax or payment is proportional only when $G_X = G_{X-T}$ such that $\pi_{MT} = 1$. Kakwani (1977) showed that this proportional relationship is valid only when the average payment rate (g) is $\frac{1}{2}$. This is because as Kakwani (1977) demonstrated, $G_X - G_{X-T} \equiv (g / 1 - g) K$. Musgrave and Thin's representation is also not invariant to scaling payment rates by a constant positive scalar. For example when all payments are increased by

²⁵ The conclusion on progressivity/regressivity based on comparing the pre-and post-payment Gini index under the Musgrave and Thin model can only be unambiguous when the financing mechanism is regressive. In the case of a progressive financing mechanism, the conclusion could be ambiguous. This is based on insights provided by the decomposition analysis of the redistributive effects. To understand this we have to note that inequality in post-payment income is affected by the tax or payment rate and therefore by unequal treatment of equals and what Abu-Zaineh *et al.* (2009) would call “improper treatment of unequals”. As we shall see later, the difference between the Gini index of pre-payment income (G_X) and post-payment income (G_N) measures redistribution associated with health care payments. This can further be decomposed into the vertical effect (V), horizontal effect (H) and re-ranking (R) such that $G_X - G_N = V - H - R$. By definition $H, R \geq 0$. From this decomposition it is easy to see that when $V > 0$ and $(H + R) > V$, inequality in the post-payment income has increased relative to that in the pre-payment income. By Musgrave and Thin's interpretation, such a financing mechanism is regressive when it is in fact progressive (a consequence of $V > 0$).

% progressivity increases. Kakwani noted that this is not consistent with the definition of progressivity (i.e. a relative measure as opposed to an absolute measure). A measure of progressivity should not capture this effect because it measures deviations from proportionality (Kakwani, 1977). This has limited the use of the Musgrave and Thin methodology in modern day empirical applications.

The Suits index (Suits, 1977) of progressivity on the other hand compares the Lorenz curve of the pre-payment income with the *relative* concentration curve of payment. Basically, if we denote the *relative* concentration index of payment as H_{pay} , Suits index of progressivity (π_s) can be defined as:

$$\pi_s = H_{\text{pay}} \quad (4.4)$$

Payments are progressive only when $\pi_s > 0$. Theoretically $0 \leq \pi_s \leq 1$. Suits index has been criticised as being sensitive to the choice of some weighting factor that is based on the slope of the Lorenz curve (Formby *et al.*, 1981). It assigns greater weights to deviations from proportionality at higher income groups than departures from proportionality at lower income groups (Formby *et al.*, 1981, van Doorslaer and Wagstaff, 1993) when it is compared with the Kakwani index for instance²⁶ (van Doorslaer and Wagstaff, 1993).

While the Suits index was being developed, Kakwani (1977) proposed a method (and an index) for assessing progressivity. The Kakwani index compares the concentration curve of payments with the Lorenz curve of pre-payment income. It is twice the area between these curves. In empirical analysis it is computed from the associated indices as:

$$\pi_K = C_T - G_X \quad (4.5)$$

where C_T is the concentration index of payments (obtained from the concentration curve) and G_X is the Gini index of pre-payment income (obtained from the Lorenz curve). A progressive (regressive) payment occurs when $C_T > G_X$ ($C_T < G_X$) while a proportional payment occurs

²⁶ The difference between Suit's index and Kakwani index as noted by Formby *et al.* (1981) becomes marked when comparing progressivity over time. While Suit and Kakwani indices respond similarly to changes in the distribution of tax or payment, they may diverge when income distribution changes (Formby *et al.*, 1981).

when $C_T = G_X$. Equivalently $\pi_K > 0$ indicates a progressive payment, $\pi_K < 0$ a regressive payment and $\pi_K = 0$ a proportional payment.

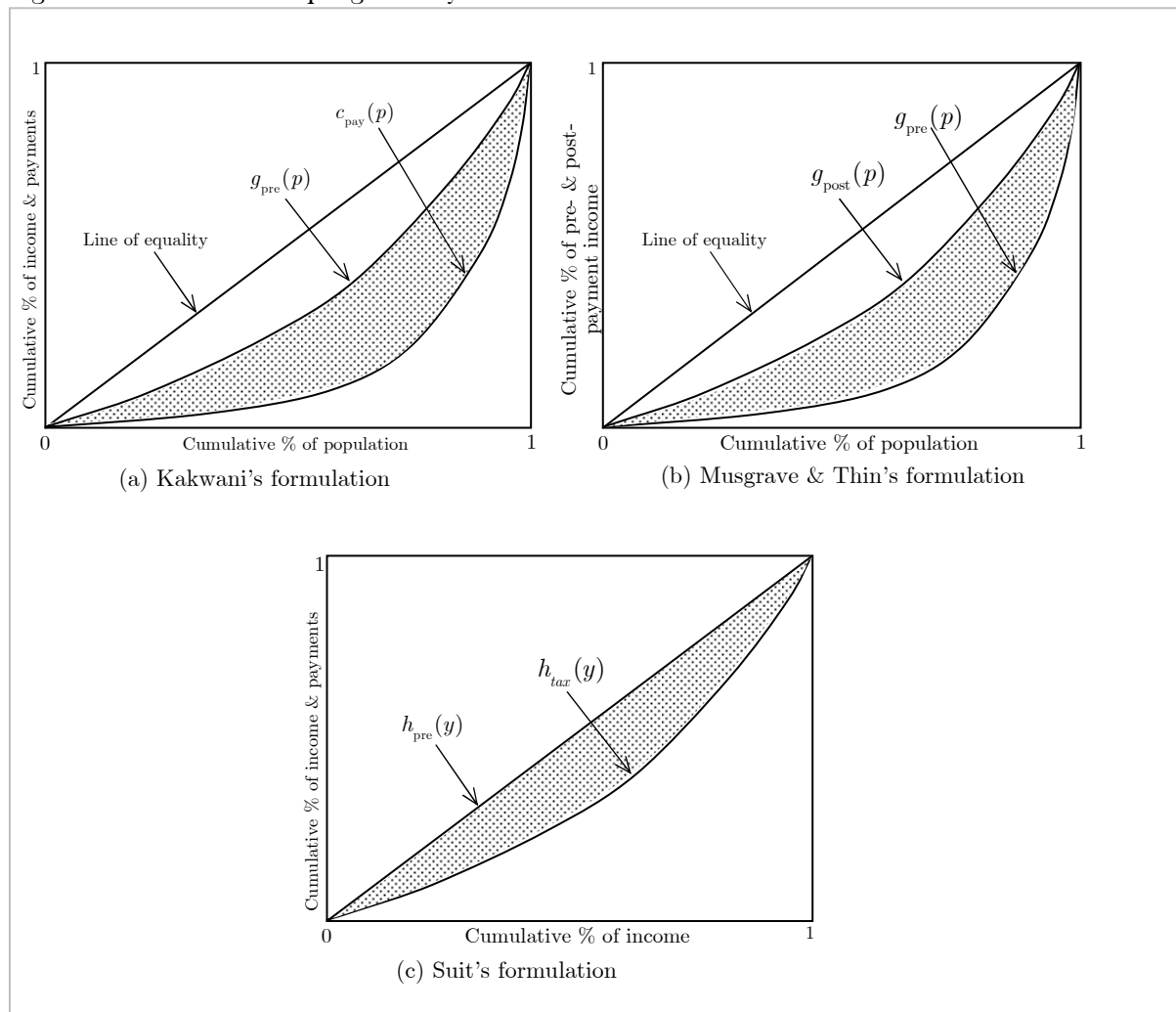
Thus while different methods have been developed, the Kakwani index has become widely used in applied research in health economics and public economics (O'Donnell *et al.*, 2008b). It has an interesting property - it is additively separable (Kakwani, 1977 p.74). This property makes it easier to obtain the overall progressivity of health care financing when Kakwani indices of each financing mechanism are known and the relative contributions of each mechanism to overall health care financing. In this case overall progressivity is obtained as:

$$\pi_K^{overall} = \sum_i w_i \pi_K^i. \quad (4.6)$$

Where π_K^i is the Kakwani index of health care financing mechanism i ; and w_i is the relative share of health care financing mechanism i (e.g. general taxes) in total health care financing. The shares are usually obtained from an external source (e.g. the National Health Accounts) (O'Donnell *et al.*, 2008b).

The various methods for assessing progressivity are summarised in Figure 4.1 using Lorenz and concentration curves. In Figure 4.1(a), the Kakwani index is obtained by comparing the Lorenz curve of income distribution $g_{pre}(p)$ with the concentration curve of tax or health care payment $c_{pay}(p)$. In Figure 4.1(b) Musgrave and Thin's index of progressivity is obtained by comparing the Lorenz curve of pre-payment $g_{pre}(p)$ and post payment $g_{post}(p)$ incomes. Here 'progressivity' occurs when there is less inequality in the post-payment income distribution than there is in the pre-payment income. Figure 4.1(c) shows the case of Suits index of progressivity. Here progressivity is assessed by comparing the relative concentration curve of pre-payment income $h_{pre}(y)$ with that of the health care payment or tax $h_{pay}(y)$. The relative concentration curve of pre-payment income $h_{pre}(y)$ plots the cumulative proportion of pre-payment income against itself. This is a benchmark curve against which progressivity is assessed. By construct, it is an analogue of the line of equality in Figure 4.1(a).

Figure 4.1: Measures of progressivity illustrated



Source: Adapted from Wagstaff and van Doorslaer (1992)

The relative concentration curve of tax or health care payment $h_{\text{pay}}(y)$ plots the cumulative proportion of pre-payment income against the cumulative proportion of tax or health care payments. In all cases, these are examples of a progressive financing mechanism and the extent of progressivity is depicted as the shaded area.

While it may be relatively easier to judge progressivity of any health care payment mechanism or system from the value of π , the curves (i.e. the Lorenz and the concentration curves) may intersect or cross each other. When this occurs, the value of the index provides an ambiguous conclusion. In respect of this, statistical dominance is used to ascertain progressivity of the mechanism across the entire ATP distribution (Bishop *et al.*, 1994, Davidson and Duclos, 1997, Sahn *et al.*, 2000, Duclos and Araar, 2006). Such dominance tests or graphical

representation of the concentration and Lorenz curves provide richer interpretation²⁷ of progressivity than does the summary index π .

One of the commonly used tests of dominance is the multiple comparison approach dominance testing procedure following Dardanoni and Forcina (1999). This has been adopted in studies in Egypt, and some selected Asian and European countries (Wagstaff *et al.*, 1999b, O'Donnell *et al.*, 2008a, O'Donnell *et al.*, 2008b). Here, curve A dominates curve B “if there is at least one quantile point at which curve A lies significantly above curve B and there is no quantile point at which curve B lies above curve A ” (O'Donnell *et al.*, 2008b p. 88). This may be compared using 19 evenly spaced quantiles from 0.05 to 0.95 (Sahn *et al.*, 2000). Statistical dominance tests in the case of inequality have been the subject of interest from economists including Bishop *et al.* (1994) and Davidson and Duclos (1997) who derive appropriate variance-covariance matrices, especially for the case of joint dependence of concentration curves and the Lorenz curve. This is relevant because most studies rely on the same data set to generate Lorenz and concentration curves. Dominance tests that ignore such dependence are less reliable (Davidson and Duclos, 1997). In empirical application most studies provide formal indices and complement these with some formal tests for statistical dominance (see for example Wagstaff *et al.*, 1999b, Abu-Zaineh *et al.*, 2008, O'Donnell *et al.*, 2008a). Other studies use only statistical dominance to assess progressivity (see Klavus (2001) for instance who used only the dominance criterion in assessing progressivity in the case of the Finnish health care financing system.)

4.2.1 Computing the Gini and the concentration indices

The Gini index is a useful measure of inequality in a distribution (usually income). It is defined as twice the area between the line of equality (45-degree line) and the Lorenz curve. The concentration index on the other hand is defined as twice the area between the concentration curve and the line of equality (Lambert, 2001, Duclos and Araar, 2006, O'Donnell *et al.*, 2008b). A Gini index of zero implies equality in the distribution of income while an index value of one implies perfect inequality. Concentration indices are bounded between negative 1

²⁷ For instance a financing mechanism could be regressive at lower levels of income but progressive at higher levels, or vice versa. This situation can result in a proportional relationship (e.g. $\pi_K = 0$) when the summary index is used. This will occur when the progressive part exactly offsets the regressive part.

and positive 1, a negative value meaning that the concentration curve lies above the line of equality, a positive value that the line of equality lies above the concentration curve²⁸.

Various formulae can be used to obtain the Gini and concentration indices (see Duclos and Araar, 2006, O'Donnell *et al.*, 2008b). Generally, we may define the Gini index for a continuous income distribution (y) as:

$$G = 1 - 2 \int_0^1 L_y(p) dp \quad (4.7)$$

Or equivalently for a discrete distribution as

$$G = \frac{2}{N\mu} \sum_{i=1}^n y_i r_i - 1 - \frac{1}{N} \quad (4.8)$$

where $L_y(p)$ is the Lorenz curve co-ordinate of income at percentile p in the distribution of income y , μ is the mean income, $r_i = i / N$ is the fractional rank of individual i in the income distribution, and N is the total number of observations. Concentration indices are analogously defined by replacing $L_y(p)$ with the concentration curve co-ordinates and μ with the mean of the variable of interest.

For computational simplicity either index can be defined using the ‘convenient covariance’ formulation (i.e. in terms of the covariance between the relevant variable and the rank of income) (Kakwani, 1980, Lerman and Yitzhaki, 1989).

$$C = \frac{2}{\mu} \text{cov}(y, r) \quad (4.9)$$

where C could be the Gini or the concentration index and y could be income (for the Gini index) or health care payments (for the concentration index).

Kakwani *et al* (1997) have used the general relationship between covariance and regression analysis to estimate concentration indices and the Gini index using the convenient regression approach as

²⁸ It should also be noted that the concentration index is a summary measure in which a positive (negative) value may not immediately mean that the concentration curve lies below (above) the line of equality. This is because these two lines may cross each other. The summary statistic only indicates which of the two effects dominates.

$$2\sigma_r^2(y_i / \mu) = \alpha + \beta r_i + \varepsilon_i \quad (4.10)$$

where y_i is again the value of the variable of interest for individual i , μ is the average of the variable of interest (i.e. income for Gini index), r_i is the weighted fractional rank of households, σ_r^2 is the variance of the fractional rank, β (the OLS estimate) is the parameter of interest - (the Gini or the concentration index). The weighted fractional rank used in equation (4.10) comes from the work of Lerman and Yitzhaki (1989) and it is computed as

$$r_i = \sum_{j=0}^{i-1} w_j + 0.5w_i \text{ with } w_0 = 0 \quad (4.11)$$

where w_i is the relative sample weight (i.e. scaled to sum up to 1) and observations are sorted in ascending order of income.

The concentration index obtained by equation (4.10) can be interpreted as the slope of a line passing through the heads of a parade of people, ranked by income, where each individual's height is proportional to the value of his/her variable of interest that is expressed as a fraction of its mean (O'Donnell *et al.*, 2008b). This method has been widely used in empirical analyses of equity in health care financing (see for example O'Donnell *et al.*, 2008a, Akazili, 2010).

Analogously we can apply (4.10) to estimate Kakwani index as defined in (4.5). This is equivalent to

$$2\sigma_r^2((z_i / \hat{\mu}_z) - (x_i / \hat{\mu}_x)) = \vartheta + \pi r_i + u_i \quad (4.12)$$

where π is the Kakwani index, z_i is the respective health care payment (e.g. private health insurance) of household i , x_i is the per adult equivalent or per capita income of household i and $\hat{\mu}_z$ and $\hat{\mu}_x$ are their respective estimated averages. The standard error of the Kakwani index may be obtained directly from (4.12) (see O'Donnell *et al.*, 2008b) or with some little transformation.

4.3 EMPIRICAL REVIEW OF PROGRESSIVITY IN HEALTH CARE FINANCING

Over the last decade and a half there have been a number of empirical studies showing the nature of the relationship that exists, within and across countries, between health care payment

and some definition of ability to pay. This is because, as noted earlier, equity in health care financing, just like taxes, is assessed by comparing payments with some measure of ability-to-pay. The usefulness of measuring and assessing progressivity extends beyond a quantification exercise. An important question to ask is: why are we interested in assessing progressivity in any health care financing system? As O'Donnell *et al.* (2008b p.187) note, the "identification of the nature of the empirical relationship and quantification of the degree of any progressivity or regressivity is of interest, not only from a wide range of equity perspectives, but also for macroeconomic and political analyses of the health care system." This could include issues of resource allocation, efficiency and improving health and productivity.

From the tax literature, as indicated above, while several methods may be used to assess progressivity (Lambert, 2001, Duclos and Araar, 2006), most studies that assess progressivity in health care financing have used either payment shares across quantiles of per capita or per adult equivalent income (Ugá and Santos, 2007), progressivity dominance (Klavus, 2001), Suits index of progressivity (Wagstaff and van Doorslaer, 1997) or the Kakwani index of progressivity (Wagstaff and van Doorslaer, 1997, Wagstaff *et al.*, 1999b). The measure of ability-to-pay (or income) used in empirical studies differs across developing and developed countries. In most developed countries per capita or per adult equivalent income is used while in developing countries per capita or per adult household consumption/expenditure is used as a proxy for household income. For instance Wagstaff *et al.* (1999b) in their study including 12 European countries used equivalent gross household income while O'Donnell *et al.* (2008a) in their study of 13 Asian countries used equivalent household consumption expenditure in all but one country (Japan) where income was the only available measure. Generally the choice between income and consumption expenditure is motivated by data constraints, especially by the reliability of income data in developing countries. These often lack well organised labour markets as so many work in the informal sector (Auriol and Warlters, 2005). Also there are multiple sources of income that exhibit a high degree of variability over time (O'Donnell *et al.*, 2008b).

An assessment of overall progressivity in health care financing usually involves considering not only payments that are made directly for health care or that are earmarked for health. It includes general tax (both direct and indirect taxes), social insurance, private insurance, and out-of-pocket payments. Where there are earmarked taxes or levies that are used to finance

health care, like the case of Ghana and Egypt (O'Donnell *et al.*, 2008b, Akazili, 2010), they are also considered.

The statutory incidence (i.e. the entity that bears the legal liability of the tax or payment) of any health care payment, like taxes, may differ from the economic incidence (the entity that bears the final burden of the tax or payment) (Younger *et al.*, 1999). This is generally the case when there are possibilities for some form of tax shifting. The exact extent to which taxes can be shifted is largely determined by the interplay between demand and supply forces. For example, the extent to which manufacturing firms can shift the burden of corporate taxes forward on to consumers depends on the prevailing market conditions (Musgrave *et al.*, 1951). Because there are likely to be differences between the statutory incidence and the economic incidence of taxes, empirical studies have to make assumptions to guide the allocation of tax and payment burdens (Wagstaff *et al.*, 1999b, Younger *et al.*, 1999, O'Donnell *et al.*, 2008a, Akazili, 2010). The most common assumptions made on the economic incidence of health care related payments include that personal income and property taxes are borne by the legal tax payer; corporate taxes borne by labour, shareholders and/or consumers; sales and excise taxes including value-added-tax are borne by consumers; employee and employer social and private health insurance contributions are borne by the employees; individual private health insurance contributions by the consumer; and the burden of out-of-pocket payments rests on the consumer (O'Donnell *et al.*, 2008b). All the studies reviewed have used these incidence assumptions in allocating the burden associated with health care payments. Based on these assumptions various different results have emerged in the literature.

Out-of-pocket payment is almost consistently a regressive or roughly proportional financing mechanism across both developed and developing countries. As shown in Table 4.1 these studies mainly used the Kakwani index to assess progressivity and the results show a general regressivity trend (as evidenced in the negative value of the index). This is also true of informal health care payments (Szende and Culyer, 2006) in Hungary. As shown in Table 4.1 except for some countries such as Bangladesh, Indonesia, Philippines (O'Donnell *et al.*, 2008a) and Egypt (O'Donnell *et al.*, 2008b), out-of-pocket payments impose either a higher burden on the poorer than the richer segments of the population or a proportional burden. Some countries such as Colombia, Guinea, Sweden, Switzerland, France and the United States have recorded Kakwani indices greater than 0.3 in absolute value. In countries where out-of-pocket payments tend to be progressive, the extent of progressivity is very small. In fact applying the statistical

test of dominance shows that they may be proportional. Though it is generally recognised that out-of-pocket payments are the most inequitable and most inefficient means of financing health care (Knaul *et al.*, 2006), health care financing in many countries, and particularly in poorer countries are dominated by direct out-of-pocket payments. In these countries, such payments account for about 45% of total health care financing while it is approximately 23% of total health expenditure globally (Lu *et al.*, 2009). In some countries out-of-pocket payments can exceed 80% of total health care financing (Knaul *et al.*, 2006, Lu *et al.*, 2009, Ataguba, 2012b).

In the Philippines, though there is some form of universal coverage, the progressive pattern is attributed to the inability of the poor to access health care meaning that they do not make payments (O'Donnell *et al.*, 2008a). In fact this is also the case in Nepal where over 80% of total health care financing is out-of-pocket payment although here this turns out to be a mildly progressive financing mechanism (O'Donnell *et al.*, 2008a). Few countries record a Kakwani index for out-of-pocket payments that is greater than 0.08. Generally countries that rely heavily on out-of-pocket payments are more likely to show greater regressive patterns. In most African countries regressivity is associated with the absence of financial protection and over reliance on out-of-pocket payments (Cissé *et al.*, 2007, Akazili, 2010) while in Europe the fact that insurance provides partial coverage in terms of services covered, is the main driver of regressive results. In Ireland, Switzerland and Netherlands for instance regressivity of out-of-pocket payment is attributed to partial insurance coverage and payments associated with the use of private general practitioners (Wagstaff and van Doorslaer, 1992).

Indirect taxes (which are also summarised in Table 4.1) indicate, in most cases and especially in high income countries, a regressive pattern. The share of tax payments in income is a decreasing function of ability to pay. These indirect taxes are a combination of different types of taxes, including sales taxes, excise taxes and value-added tax, although most of these studies simply lump them together as indirect taxes. They are progressive in some countries in Asia and Africa such as Bangladesh, China, Egypt, Ghana, Hong Kong, Indonesia, Malaysia, Nepal, Philippines, Punjab state in India, Spain (in 1980), Taiwan and Thailand (Wagstaff *et al.*, 1999b, O'Donnell *et al.*, 2008a, O'Donnell *et al.*, 2008b, Akazili, 2010) but regressive in all European countries. The reason at least in part for the seeming progressivity of these taxes in the former countries of Africa and Asia rests on the extent of informality in their economies as countries with large informal sectors tend to have a progressive pattern of indirect taxes. In turn this may be attributed to the lower likelihood of evading these taxes in formalised

economies compared to developing economies (Emran and Stiglitz, 2005). In fact tax policies and structures vary across developing and developed countries (Gordon and Li, 2009), inevitably being more sophisticated in the latter. The progressivity of indirect taxes in the case of the developing country of Ghana is due to the progressivity of both import duty and value-added tax which more than offset the regressivity of the fuel/kerosene levy. The progressivity of value-added tax is attributed to tax exemption on commodities consumed mainly by the poor (Akazili, 2010). In Spain while indirect taxes were progressive in 1980, they became regressive by 1990 due to the introduction of a value-added tax system in 1986 (Wagstaff *et al.*, 1999b).

Direct taxes in contrast to indirect taxes are consistently progressive across countries. The progressivity of direct taxes is expected *a priori* because they are progressively structured in nearly all the countries. Apart from these general taxes, some countries such as Egypt and Ghana rely on some form of earmarked taxes. Egypt imposes an earmarked health tax on cigarette sales that accounts for about 3% of total health expenditure which has been reported as being statistically proportional (O'Donnell *et al.*, 2008b). Also in Ghana 2.5% of value-added tax (VAT) rate is earmarked for health as a National Health Insurance Levy. While this accounts for less than 3% of total health expenditure, it has nonetheless been reported as mildly progressive or proportional (Kakwani index = 0.049) (Akazili, 2010).

Combining the direct and indirect taxes, general taxes are consistently progressive in all countries reviewed (see Table 4.1). The only exception to this was Denmark in 1981 but by 1987 general taxes became progressive (Wagstaff *et al.*, 1999b). The progressivity of general taxes is attributable to the share of direct taxes in total tax receipts and the general progressivity of direct taxes (O'Donnell *et al.*, 2008a). As a general rule then general taxes are progressive in countries where direct taxes are progressive and where the share of direct taxes in general taxes is high compared to regressive indirect taxes.

In most countries the contribution of social insurance, as summarised in Table 4.1, is generally a progressive financing mechanism. This form of insurance varies across countries in terms of the contribution rates, enrolment criteria and breadth of coverage. Some countries like Switzerland and Netherlands have more than one form of social health insurance covering different categories of people (Wagstaff and van Doorslaer, 1997, Bilger, 2008). In Taiwan, Japan and the Korean republic a form of universal coverage exists, funded mainly through social insurance contributions of employees and employers with some subsidies from the state.

Thailand used to operate a form of social insurance for civil servants and formal private sector employees which was later extended in 2001 to the rest of the population through additional tax funding (Limwattananon *et al.*, 2007, O'Donnell *et al.*, 2008a). The question of who in general bears the burden of social insurance will therefore depend on the category of the population covered, the extent of reliance on this form of financing and how the rates are structured. From the review of studies that investigate the progressivity of social insurance, countries that rely to a sizeable extent on social insurance include France, Japan, Taiwan, Netherlands, Germany, and republic of Korea. Among these countries, apart from France, social insurance contributions were mildly regressive in Japan, Taiwan and Germany (Wagstaff *et al.*, 1999b, O'Donnell *et al.*, 2008a) and more clearly regressive in Netherlands and the republic of Korea (Wagstaff *et al.*, 1999b, O'Donnell *et al.*, 2008a). The regressivity of social insurance contributions in Korea has reduced in recent years because the upper earning cap on contributions was abolished in 2002 (O'Donnell *et al.*, 2008a). Therefore the more a country relies on this financing mechanism, the more likely there is to be a regressive relationship between contributions and ability-to-pay, largely because the poorer segments of the population tend to buy into the scheme. In Japan and Taiwan the mildly regressive to proportional relationship is attributable to the nearly constant average contribution rates across earnings (O'Donnell *et al.*, 2008a). In the Netherlands the regressive pattern in 1987 and 1992 is attributed to non-involvement of the better-off in all or part of the Dutch social insurance system and in Germany this was because individuals who earn above a specified threshold can opt out from the social scheme (Wagstaff *et al.*, 1999b). In other Asian countries where social insurance was progressive, this is largely linked to partial coverage where only the richer segment of the population contribute to the social insurance system (O'Donnell *et al.*, 2008a). In general while social health insurance has a long history in Europe and some parts of Asia, it is still under developed and is a very uncommon financing mechanism in Africa.

Private health insurance coverage, with few exceptions, is relatively low in the developing countries of Asia and Africa. The incidence of private insurance contribution could also be either progressive or regressive depending largely on the type of private insurance that exists in the country. While there may be other categories that overlap, three broad forms of private insurance can be distinguished. One form of private cover especially in developed countries such as France and Denmark is that which provides cover against co-payments charged at public sector facilities. Another form occurs in countries such as Spain, Portugal, the UK, and Italy where private cover is taken out to supplement other cover provided by the government.

The third form of private insurance that exists in countries like the United States, formerly in Switzerland, the Netherlands and Germany is taken out as the sole cover (Wagstaff *et al.*, 1999b). Generally in countries such as Portugal, Italy and the UK where private cover is supplementary, progressive patterns are observed. This is because those who can afford premiums are the richer segments of the population. In Switzerland private insurance is bought by a large proportion of the population and was a very regressive financing mechanism between 1982 and 1992. The same is true of the United States. In fact as presented in Table 4.1 these two countries record the highest absolute value of the Kakwani index. In Germany, private health insurance is progressive because it is mainly the rich who opt out from social insurance that purchase private cover (Wagstaff *et al.*, 1999b). In the case of Palestine, though private health insurance accounts for about 10% of total health care financing, its catchment population is the richer groups. It is thus not surprising that it appears as progressive. It is driven largely by exclusion of the poor through the fact that the premiums are about three times those of the Government Health Insurance (GHI), a form of social health insurance (Abu-Zaineh *et al.*, 2008).

The overall progressivity or regressivity of a country's total health care financing system is a function of the relative shares of each financing component and their relative progressivity. Generally, as summarised in Table 4.1, most health systems are either mildly regressive or nearly proportional. The regressivity observed in the overall health care financing systems of most countries is attributed to reliance on regressive finance mechanisms. Countries with progressive health systems rely less on regressive mechanisms and finance health care mainly through progressive taxes.

In summary, the extent of progressivity in health care financing varies from country to country. Indirect taxes are generally seen as regressive health care financing mechanisms while direct taxes are progressive. There are differences across countries relating to the progressivity of private and social health insurance and out-of-pocket payments. Generally, the review shows that the extent of progressivity observed across countries is not necessarily correlated with GDP per capita. This conclusion is drawn from the summary provided in Table 4.1 where developed countries with high GDP per capita such as the United States, Sweden, Switzerland and Portugal are shown to have regressive health care financing systems while poorer countries like Sri Lanka, Thailand and Bangladesh are seen to have relatively progressive financing systems. Again the ratio of a country's total health spending to GDP may not be a guide to its

progressivity. For instance in Brazil public health funding alone accounts for about 3.8% of GDP (Ugá and Santos, 2007) yet overall the financing system was almost regressive. However Malaysia that spends about 3.8% of GDP on health (public and private) shows a very progressive financing system (Yu *et al.*, 2008, 2011). The overall determinant of a progressive system is therefore the extent to which the poor are shielded and how the health system is financed and not necessarily the level of a country's income. Over reliance on regressive mechanisms will lead to a regressive total health care financing system.

4.4 DECOMPOSING REDISTRIBUTIVE EFFECT OF HEALTH CARE PAYMENTS

Like taxes, health care payments may treat equals unequally or unequals unequally and may also cause the reranking of households from richer to poorer ranks or vice versa (Aronson *et al.*, 1994, Aronson and Lambert, 1994, van de Ven *et al.*, 2001). While the assessment of progressivity provides a way of indicating the extent to which health care payments deviate from proportionality, it does not tell us how post payment incomes are unequally affected and what drives such distribution. If therefore concerns about how health care payments affect income are important, their effect on income distribution must also be seen as relevant (Wagstaff and van Doorslaer, 1997). Specifically, the redistributive effect quantifies how much more unequal (or equal) health care payments make the distribution of income (Wagstaff and van Doorslaer, 2001).

For simplicity, the easiest way to measure the redistributive effect of taxes or health care payments is by comparing inequality in income pre and post payment (Reynolds and Smolensky, 1977). This is now the standard approach in redistribution studies. Technically, indices of income inequality (e.g. the Gini index or the Atkinson class indices) are often used (see Reynolds and Smolensky, 1977, Lambert, 2001) to assess redistribution.

Redistribution is defined, following Reynolds and Smolensky (1977)²⁹ as:

$$RE = G_X - G_{X-T} \quad (4.13)$$

²⁹ Though Musgrave and Thin (1948) originally proposed, they did not call it a measure of redistributive effect but a measure of progressivity.

where G_X and G_{X-T} are the pre- and post-payment Gini coefficients respectively³⁰, X is the pre-payment income gross of health care payments and direct taxes, and T represents health care payment. A pro-poor (pro-rich) redistribution is that which redistributes income from the rich (poor) to the poor (rich). A proportional redistribution leaves the post-payment income distribution unaltered. Aronson *et al.* (1994) hereafter referred to as *A/L* formalised Reynolds and Smolensky's (1977) formulation and showed that (4.13) can be decomposed as follows:

$$RE = V - H - R \quad (4.14)$$

where each of the components of the right-hand-side in (4.14) measures a different dimension of equity (or inequity) in the same unit of measurement (Bilger, 2008). Specifically, V denotes vertical redistribution which according to Bilger (2008) measures the extent to which a financing mechanism or system is pro-poor or pro-rich. It also measures the tendency of a financing system to 'compress' the distribution of net incomes (Duclos and Araar, 2006). H captures the 'classical' definition of horizontal inequity; and R the extent of reranking defined in the 'classical sense' as a change in the order of income distribution. On occasion, H (horizontal inequity) and R (re-ranking) have sometimes been taken to mean the same (see Wagstaff and van Doorslaer, 1997, Duclos and Araar, 2006)³¹ but Aronson and Lambert (1994) have shown that these are two distinct even if related concepts (see also Wagstaff and van Doorslaer, 2001). Re-ranking relates to the "improper treatment of unequals" (Abu-Zaineh *et al.*, 2009) or how post-payment income ranks do not correlate with pre-payment income ranks while horizontal inequity is related to treatment of equals. However it is to be noted that it is possible for horizontal inequality to induce reranking (Bilger, 2008). For example "a health insurance scheme with premiums fixed according to risk causes both horizontal inequality and reranking" (Bilger, 2008 p.1584). However, it is possible for reranking to also occur in the presence of horizontal equity. This occurs "when marginal financing exceeds unity in some parts of the income distribution" (Aronson *et al.*, 1994, Bilger, 2008 p.1584). This can happen for instance where individuals who earning income up to a

³⁰ In a later formulation Duclos *et al.* (2003) use the Atkinson type inequality measures. The redistributive effect is still measured as the difference between inequality in pre- and post-payment income.

³¹ Horizontal inequity is the equal treatment of *equals* while reranking relates to treatment of *unequals* (Wagstaff and van Doorslaer, 1997). Reranking can be divided into three types - within-group, entire group, and partial or inter-group reranking (see Urban and Lambert, 2008). This distinction is important in trying to understand how many authors have used re-ranking.

certain threshold receive ‘free’ health care services (i.e. there is no payment at the point of consumption). It can then be that individuals who are towards the bottom of the next higher income threshold may face payments that are likely to rerank them.

This redistributive effect is of policy importance when first we are interested in the amount of resources households may be left with to purchase other basic necessities after making health care payments (either directly or through taxes or insurance) and second for other equity reasons. Such analysis, apart from showing explicitly the extent to which post payment incomes are affected, also reveals the redistributive consequences of each financing mechanism. This can inform policy makers who are interested in equity with options to reform health care financing systems. At the same time this can reveal the constituents of redistribution.

4.4.1 Developments in assessing redistributive effects

Early developments of indices for assessing progressivity include that initially proposed by Musgrave and Thin (1948) and explicitly shown by Kakwani (1977) to be defined as:

$$G_X - G_{X-T} = (g / g - 1) \pi_K \quad (4.15)$$

where g is the average payment or tax rate (i.e. the share of payment in income) and the rest remain as defined earlier. Kakwani (1977) was the first to show explicitly that payment rates bear a positive relationship with redistribution. The higher the payment rate the lower the inequality in post-payment income and by definition the higher the redistributive effect of payment. Therefore in moving from a pre- to post-payment income distribution, higher values of g reduce inequality in post-payment income and increase the redistributive effect of payments. Similarly from this formulation, a more progressive tax or health care payment reduces inequality in post-payment incomes. These results can be deduced from the elasticity of post-payment income inequality with respect to average payment rate (g):

$$\eta_g = (g \cdot \pi_K) / ((1 - g)^2 G_{X-T}) \quad (4.16)$$

and the elasticity of post-payment income inequality with respect to π_K given as:

$$\eta_{\pi_K} = -(g \cdot \pi_K) / ((1 - g) G_{X-T}) \quad (4.17)$$

The ratio of these elasticities is greater than unity in absolute terms ($|\eta_{\pi_K} / \eta_g| > 1$) implying that redistribution is “more sensitive to the average tax rate than to tax progressivity” (Kakwani, 1977 p.73). Therefore a system that is progressive could produce a greater pro-poor redistribution when the average tax or payment rate is increased than when progressivity is increased.

In 1984 Kakwani amplified the formulation in (4.15) to account for the likelihood of reranking (Kakwani, 1984). This later amplification however does not account for the presence of horizontal inequity. It simply assumes horizontal equity. We may write Kakwani’s (1984) formulation as follows:

$$RE \equiv G_X - G_{X-T} = \underbrace{\left(\frac{g}{1-g} \right) \pi_{\bar{K}}}_{V^K} - \underbrace{R}_{R^K} \quad (4.18)$$

Where g is the average payment rate, $\pi_{\bar{K}}$ is the Kakwani index of progressivity and which implicitly assumes the absence of horizontal inequity (Aronson *et al.*, 1994). Here it is again assumed that the higher the average payment or tax rate, the higher the redistributive effect of payment. Reranking in this framework increases inequality in post-payment income relative to inequality in pre-payment income. Similarly, as noted earlier, a progressive financing system (i.e. $\pi_{\bar{K}} > 0$) reduces inequality in post-payment income.

The apparent absence of horizontal inequity implicit in earlier formulation, for example Kakwani’s (1984), will not always be plausible as households with the same pre-payment income may end up contributing different amounts for health care. This, as noted earlier, could result from different scenarios ranging from households with multiple sources of income to choosing different insurance packages. Some of these treatments may arise as a result of choices made by individuals and as such may not fit the classical definition of “horizontal inequity” (Wagstaff *et al.*, 1999a). These choices may result for differences in tastes or in risk aversion. Therefore caution is needed in labelling any horizontal effect as inequitable. A good assessment of this will require an understanding of the health care financing system of the country and also how payments are made. One major challenge with assessing horizontal inequity is the difficulty in defining equals since its measurement is clearly sensitive to the way in which equals are categorised (Aronson *et al.*, 1994, Urban and Lambert, 2008). Because of

these difficulties and subsequent confusion, some may argue that horizontal inequity is not a goal that is worth pursuing. If however one were to argue simply that all human beings are worth the same, inequity becomes a purely horizontal affair.

In many instances the unequal treatment of ‘comparable’ equals may generate some resentment and a sense of insecurity which might in turn lead to social unrest. This has been corroborated in socio-psychological studies relating discrimination, alienation and exclusion to social cohesion and welfare (see Duclos *et al.*, 2003). The same is also true of deprivation seen in relative terms (Runciman, 1966) as by definition individuals then evaluate their position relative to those with similar standing. It may well be that these various debates about the appropriateness of horizontal inequity as an equity issue have underpinned the absence of horizontal inequity in earlier decompositions of redistributive effects. More likely however the simple explanation is that horizontal inequity is abhorred as a result of the common aversion to inequality that exists among comparable equals (Duclos *et al.*, 2003).

Aronson *et al.* (1994), in a classic paper, advanced the idea of the decomposition of redistribution by explicitly including horizontal inequity as a component. They defined the individual components of equation (4.14) after grouping individuals or households into groups with similar pre-payment income levels³² as follows:

$$V^{AJL} = G_X - G_{X-T}^B \quad (4.19)$$

or equivalently decomposed as

$$V^{AJL} = (g / 1 - g)\pi_{\bar{K}} \quad (4.20)$$

³² This is done because the probability of finding individuals in a survey with the same level of pre-payment income (Duclos *et al.*, 2003, Bilger, 2008) is not high. If these do exist, they are usually few. What we require is that we have individuals or households in groups with similar pre-payment incomes (or ability to pay) – equals. While in this context the choice of income intervals or bandwidth is of critical importance, such a choice does not affect the measure of $(H + R)$ per se but only their relative magnitudes. In general, H (R) increases (falls) with larger pre-payment income intervals (see Aronson *et al.*, 1994, O'Donnell *et al.*, 2008b). Also, reranking may be caused not only by differential treatments of households with similar pre-payment income but due to, for instance as Aronson *et al.* (1994) noted, marginal tax rates greater than 100%. Where R is caused only by differential treatment, knowing the individual values of H and R is of little policy importance compared to the situation where R is attributed mainly to marginal tax rates greater than unity (Wagstaff and van Doorslaer, 1997). The use of groups of similar individuals is one of the major criticisms and drawbacks of the AJL methodology.

where G_{X-T}^B is the between-group Gini coefficient for post-payment income (computed by replacing all post-payment incomes with their group means) and $\pi_{\bar{K}}$ is the Kakwani index if there were no horizontal inequity in health care payments. This is the difference between the between-group concentration index and the Gini index of pre-payment income. V^{DJA} measures regressivity or progressivity in financing given that there is no reranking and there is no horizontal inequity.

Horizontal inequity is computed as:

$$H^{AJL} = \sum_j \alpha_j G_{X-T,j} \quad (4.21)$$

where $G_{X-T,j}$ is group j specific post-payment income Gini coefficient; α_j is the product of each group's population share and its post payment income share. In short, H^{AJL} is interpreted as the weighted sum of the group specific post-payment income Gini coefficients. By construction it is non-negative and can only reduce redistribution. Hence horizontal inequity makes the distribution of post payment income more unequal than in its absence (O'Donnell *et al.*, 2008b).

Reranking is computed as:

$$R^{AJL} = G_{X-T} - C_{X-T} \quad (4.22)$$

Where C_{X-T} is the post-payment income concentration index obtained by ranking households first by their pre-payment income then within each group of equals, by their post-payment income. Also, R^{AJL} is non-negative - it can only reduce redistribution³³.

Therefore we can re-write equation (4.14) as:

$$\underbrace{G_X - G_{X-T}}_{=RE} = \underbrace{(g / 1 - g)\pi_{\bar{K}}}_{=V^{AJL}} - \underbrace{\sum_j \alpha_j G_{X-T,j}}_{=H^{AJL}} - \underbrace{G_{X-T} - C_{X-T}}_{=R^{AJL}} \quad (4.23)$$

In a similar reformulation of the AJL methodology, van de Ven *et al.* (2001) have shown that equation (4.23) is numerically equivalent to:

³³ It is not difficult to reason this since the concentration curve cumulates post-payment incomes in increasing order of the pre-payment income while the Lorenz curve cumulates post-payment incomes in increasing values of the post-payment income. Hence the concentration curve for post-payment income will never be lower than the Lorenz curve for post-payment income.

$$\underbrace{G_X - G_{X-T}}_{=RE} = \underbrace{G_X - G_{X-T}^B}_{=V^{AJL}} - \underbrace{\sum_j \alpha_j G_{X-T,j}}_{=H^{AJL}} - \underbrace{E_{X-T}}_{=R^{AJL}} \quad (4.24)$$

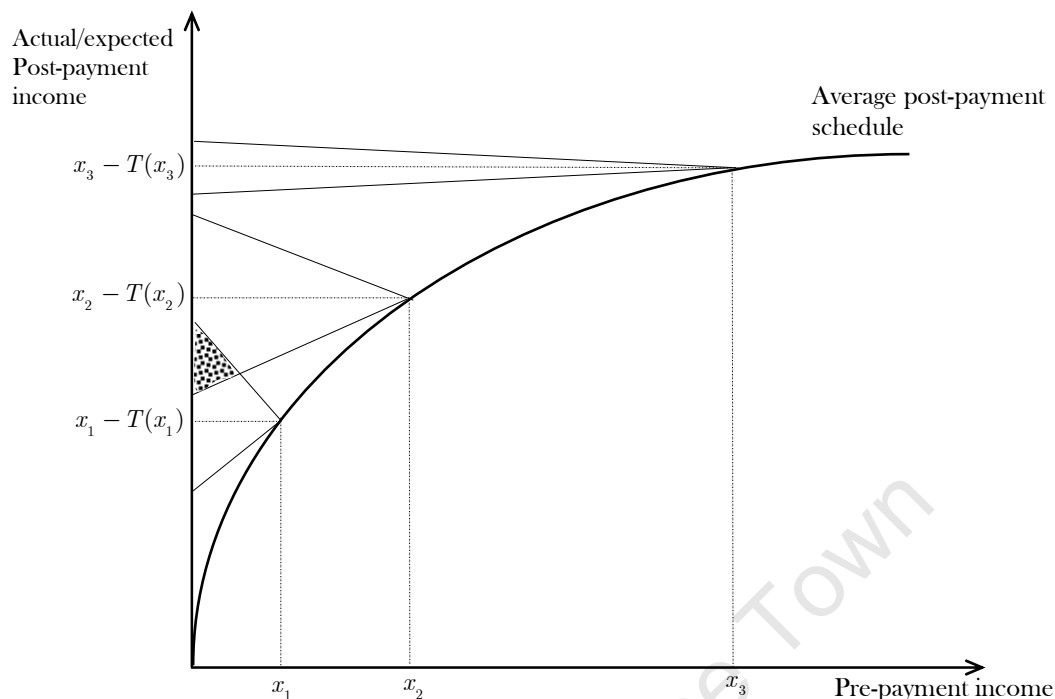
The last component E_{X-T} measures reranking and this is the residual from decomposing the Gini coefficient of post-payment income (van de Ven *et al.*, 2001, Urban and Lambert, 2008). This is claimed as *similar* to the Atkinson-Plotnick measure of reranking (Atkinson, 1979, Plotnick, 1981) without the usual normalization by $2G_{X-T}$. But as recently derived by Urban and Lambert (2008), the formulation E_{X-T} based on the AJL framework (R^{AJL}) is not necessarily an analogue of that proposed by Atkinson and Plotnick and Kakwani because it omits the within-group reranking and the re-ordering of the entire group³⁴ (see Urban and Lambert, 2008).

In summary the AJL framework decomposes the total redistributive effect into four major components viz. (a) average rate effect (g); (b) departure from proportionality or progressivity effect ($\pi_{\bar{K}}$); (c) horizontal inequity effect H ; (d) reranking effect R .

In Figure 4.2, a simplified illustration of the vertical effect, horizontal effect, and reranking are presented as adapted from Aronson *et al.* (1994). The horizontal axis plots pre-payment income while the vertical axis plots the actual/expected post-payment income. On the vertical axis, $x_i - T(x_i)$ measures the expected post-payment income for an individual with pre-payment income x_i . This is the expected post-payment income if all individuals with the same pre-payment income x_1 contributed equally (e.g. through taxes, insurance premiums) to the health system. However, there are often deviations (Aronson *et al.*, 1994) which are represented by the ‘fans’ and are attributed to disturbances $\varepsilon(x_i)$. These ‘fans’ show the extent to which actual post-payment incomes are distributed around the expected post-payment income for a given pre-payment equal. The fans provide an indication of unequal treatment of individuals with the same level of pre-payment income. Horizontal inequity occurs when $\varepsilon(x_i) \neq 0$ in any pre-payment income group (i.e. where the fans exist).

³⁴ This issue was discussed further under the Urban and Lambert (2008) framework.

Figure 4.2: A descriptive illustration of the components of redistributive effects



Source: Aronson *et al.* (1994).

When the ‘fans’ overlap, this leads to the reversal or switching of ranks. This occurs in the shaded region in Figure 4.2. This shows that some individuals with pre-payment incomes x_2 (given that $x_2 > x_1$) became worse-off than some individuals with pre-payment income x_1 as a result of health care payments. This is what is captured as differential treatments. It is important to note that the framework presented in Figure 4.2 assumes that the presence of horizontal inequity (i.e. where the fans exist) gives rise to re-ranking (Aronson *et al.*, 1994). However, as discussed earlier, this need not be the case. When the marginal tax or payment rate is greater than unity ($T'(x) > 1$), such as is the case where the average post-payment schedule in Figure 4.2 exhibits a dip, re-ranking can also occur. When this occurs, it is also intuitive to assume that overlapping ‘fans’ generate unintended effects (Aronson *et al.*, 1994 p.264). The vertical effect of redistribution is implicit in the concave schedule in Figure 4.2 (an implication of progressivity).

The AJL methodology dominated early studies on redistribution in taxes especially in developed economies. However, one of its major limitations is that it assumes that equals of pre-payment income can be obtained (van de Ven *et al.*, 2001, Bilger, 2008, O'Donnell *et al.*, 2008b) which is often not the case in practice. As a result, most often, *near-equals* are

constructed using some income bandwidths or intervals (Aronson *et al.*, 1994). The choice of this bandwidth creates some problems in the way horizontal inequity is defined. It loses its ‘classical’ definition. As noted above, the choice of the bandwidth does not affect the measure of $H + R$ but only their relative magnitudes. Nonetheless, it may well be of interest to know their relative magnitudes as these are two distinct measures with likely distinct policy implications (Duclos *et al.*, 2003). Others such as van de Ven *et al.* (2001) and Duclos *et al.* (2003) have investigated this weakness by proposing modifications to assess the redistributive effects in a tax or financing system that may obviate the choice of near equals.

van de Ven *et al.* (2001) - referred to as VCL modified equation (4.24) as follows:

$$\underbrace{G_X - G_{X-T}}_{=RE} = \underbrace{(G_X^B - G_{X-T}^B)}_{=V^{VCL}} - \underbrace{\sum_{j=1}^J (\alpha_j G_{X-T,j} - \alpha'_j G_{X,j})}_{=H^{VCL}} - \underbrace{E_{X-T}}_{=R^{VCL}} \quad (4.25)$$

where all old components in (4.25) are defined as before. The new components G_X^B , $G_{X,j}$ and α'_j are the between-group Gini coefficient of inequality in the pre-payment income, the group specific Gini coefficient of inequality in the pre-payment income, and the product of the group’s population share and its pre-payment income share respectively.

van de Ven *et al.* (2001 p.383) have also shown that (4.25) is a special case of or rather equal to equation (4.24) only when all groups contain exact pre-payment equals such that $G_{X,j} = 0$ for all j and $G_X^B = G_X$. This is because between inequality is the only source of inequality.

More recently, Urban and Lambert (2008) have shown that the decomposition proposed in (4.25) is consistent with economic theory. This is because it can be shown that (4.25) can be decomposed such that

$$G_X = G_X^B + \sum_j \alpha'_j G_{X,j} + E_X \text{ where } E_X = 0 \quad (4.26)$$

and

$$G_{X-T} = G_{X-T}^B + \sum_j \alpha_j G_{X-T,j} + E_{X-T} \quad (4.27)$$

where equations (4.26) and (4.27) represent the Gini index of pre-payment and post-payment incomes across near-equals respectively.

While in Aronson *et al.* (1994), the choice of the bandwidth is to an extent arbitrary, van de Ven *et al.* (2001) proposed a simple “spline” methodology for decomposing the redistributive effect as well as an innovative criterion for the choice of an optimal bandwidth. They note particularly two effects – averaging and appropriation effects – that may result from choosing large bandwidths. For the averaging effect, which they called an improving effect, increasing bandwidth increases the number of pre-payment ‘near-equals’ in a group, and it also “reduces the variability of the function linking group average net income to gross income” (Bilger, 2008 p.1585). When bandwidths are increased, this leads to a reduction in the number of groups to be formed. This gives rise to the appropriation effects where increasing proportions of the V and the R observed in a given sample are attributed to H (van de Ven *et al.*, 2001). Also and more importantly, the averaging and appropriation effects decrease with increasing bandwidths. This implies that there will be a bandwidth such that the combination of the two effects will minimise errors associated with accurate decomposition estimates of V , H and R . When the bandwidth is very small, there is little averaging which results in high errors in decomposition. However, when the bandwidth is increased, there will be gains in decomposition accuracy (as a result of the averaging effect in action) but a loss caused by the appropriation effect. However “the gains derived from the averaging effect diminish relative to the losses associate (sic) with the appropriation effect as w (*the bandwidth*) is increased, which implies an optimal w that minimises the decomposition error” (van de Ven *et al.*, 2001 p.388). These averaging and appropriation effects have diverging impacts on the decomposition estimates of V , H and R . van de Ven *et al.* (2001) specifically note that the appropriation effect has an ambiguous impact on H . What this means is that we need to choose the bandwidth that maximises the estimate derived from V (i.e. the bandwidth that corresponds to the maximum value of V) i.e. V_{\max}^{VCL} . From this, the value of H^{VCL} is computed accordingly as:

$$H^{VCL} = V_{\max}^{VCL} - R^{AJL} - RE \quad (4.28)$$

where, also in this case, R^{AJL} is analogous to the AJL formulation of reranking.

It is important to note that while the VCL approach of choosing V_{\max}^{VCL} is appealing, its application is limited to only the case where the financing mechanism or system is progressive (Bilger, 2008). When the financing system or mechanism is regressive, there is no maximum for V^{VCL} as the averaging and appropriation effects contribute to an increase in V^{VCL} as opposed to an alternating effect observed in a progressive system when the bandwidth is

increased. Therefore in this regard the VCL methodology is limited or of no value for empirical analysis (Bilger, 2008) in countries with regressive mechanisms.

The foregoing discussions show that while the AJL decomposition framework suffers from arbitrariness in choosing the bandwidth, the VCL approach, though consistent with economic theory (Urban and Lambert, 2008), suffers from the same drawback when the financing mechanism is regressive (see Bilger, 2008 for an example of the impact of bandwidth choice).

Urban and Lambert (2008), referred to here as UL, propose some modifications of the AJL and VCL formulations. They noted that while attempts have been made to account for near-equals group based on the AJL and the VCL formulations for instance, there still remain some fundamental representations that are omitted when decomposing the redistributive effect of a financing system. The computation of the reranking effect in the AJL and the VCL models (R^{AJL}), for instance, does not incorporate within group reranking³⁵ (R^{WG}) and entire group reranking³⁶ (R^{EG}) (see Urban and Lambert, 2008). Therefore, under the UL model, the measure of reranking is defined as $R^{AP} = R^{WG} + R^{EG} + R^{AJL}$. Similarly, UL defines V^{UL} as vertical inequity with “full vertical effect”³⁷ and H^{UL} as horizontal inequity with “type-2 horizontal effect”³⁸. Urban and Lambert (2008) note further that of the AJL, VCL and the UL

³⁵ Here ranks of households that are grouped as close-equals may switch. This is intuitive because the close equals are determined using some income bandwidth.

³⁶ This occurs when an entire group of defined close-equals change rank.

³⁷ These are referred to as full vertical effect because unlike the AJL formulation, they recognise the possibility of whole group re-ranking. Urban and Lambert (2008) in fact showed that $V^{UL} = V^{AJL} + R^{EG}$. Vertical effect here is computed as: $V^{UL} = G_X - C'_{X-T}$. Where G_X is the pre-payment income Gini index and C'_{X-T} is the concentration index of post-payment income that is obtained by deducting within each group, the group's (k) payment share (g_k) of income from each household's pre-payment income.

³⁸ Type-2 effect is attributed to the work of King (1983) and later Jenkins (1994). This involves comparing the distribution of household-by-household departures of actual post-payment income with another schedule that is counterfactually constructed to be horizontally equitable within groups of close equals. Horizontal effect H^{UL} is therefore obtained as: $H^{UL} = C_{X-T} - C'_{X-T}$. Where C_{X-T} is the post-payment concentration index when households are ranked by their pre-payment income and by post-payment income for exact pre-payment income equals. C'_{X-T} is as defined earlier.

decompositions, only that of UL takes account of all the forms of reranking ($R^{WG} + R^{EG} + R^{AJL}$). The empirical application of the UL in the assessment of equity in health care financing is limited although it has been recently applied to assess the redistributive effect of health care financing in the Occupied Palestinian Territory (Abu-Zaineh *et al.*, 2009).

In summary, the UL formulation can be written as:

$$RE = \underbrace{G_X - C'_{X-T}}_{V^{UL}} - \underbrace{C_{X-T} - C'_{X-T}}_{H^{UL}} - \underbrace{(R^{EG} + R^{WG} + R^{AJL})}_{R^{UL} = R^{AP}} \quad (4.29)$$

The various components of R^{AP} can be obtained as:

$$R^{WG} = C''_{X-T} - C_{X-T} \quad (4.30)$$

$$R^{EG} = C'''_{X-T} - C''_{X-T} \quad (4.31)$$

$$R^{AJL} = R^{BG} = G_{X-T} - C'''_{X-T} \quad (4.32)$$

Where C_{X-T} is the post-payment income concentration index when households are ranked by their pre-payment income and then by their post-payment income for exact pre-payment income equals; C''_{X-T} is the concentration index of post-payment income when households within each group are ranked by post-payment income and then each group is ranked by the mean of the pre-payment income (i.e. where there is no between group re-ranking); and C'''_{X-T} is the concentration index of post-payment income when households within each group are ranked by post-payment income and then each group is ranked by the mean of the post-payment income.

While UL representation explores the deficiencies in AJL and VCL, it has not been able to account for the effect of the choice of near-equals as opposed to actual equals. This is because it still relies on the use of income groups. Duclos *et al.* (2003 p.67) note that “estimation difficulties were deemed in the past to limit severely the practical relevance of some approaches to measuring VE [*vertical equity*] and HI [*horizontal inequity*].” On this Bilger (2008 p.1585) notes that “the drawback of not measuring horizontal inequality in the classical sense and adversely affecting the measurement of the decomposition effects remains anyway.” Because of the difference between the classical definition of horizontal inequity and reranking (see also Duclos and Araar, 2006) it is important to measure horizontal equity based on the classical definition of treating equals equally.

Duclos *et al.* (2003) (sometimes referred to here as *DJA*) therefore propose a novel methodology to assess redistributive effects based on a different evaluative framework, what they term ‘ethical’ social welfare functions. The methodology attempts to solve the arbitrariness inherent in the choice of near-equals so as to measure pure ‘classical’ horizontal inequity. In doing so near-equals are not chosen by the analyst or decision maker but equals are statistically determined using non-parametric approaches. Duclos *et al.*’s (2003) decomposition follows the traditional method proposed by Reynolds and Smolensky (1977). They also define redistribution as the difference between inequality in pre- and post- payment incomes but rather than using the traditional Gini based inequality indices, they define their change in inequality approach using some Atkinson- type inequality measures that incorporate some ‘ethical’ concerns. Specifically they use Yaari’s (1988) specification of the Atkinson class of social welfare functions $W(\varepsilon)$ (Atkinson, 1970). A form of Yaari’s (1988) social welfare formulation that is homothetic³⁹ is specified as

$$W_X(\varepsilon) = \int_0^1 U_\varepsilon(X(p)) w(p) dp \quad (4.33)$$

where $X(p)$ is the gross income of the individual with rank p in the income ladder or the pre-payment quantile income function, ε is the parameter of aversion to horizontal inequality. However the formulation in equation (4.33) suffers from the same criticisms which are levied against a purely utilitarian utility function. They are not sensitive to the distribution of utilities (see Feldstein, 1976). In order to accommodate such concern, Duclos *et al.* (2003) modify equation (4.33) by including $\omega(p; \nu)$ as shown in equation (4.34). This utility formulation in equation (4.34) is flexible to accommodate distributional concerns that are central to the assessment of equity and fairness.

$$W_X(\varepsilon, \nu) = \int_0^1 U_\varepsilon(X(p)) \omega(p; \nu) dp \quad (4.34)$$

Here $\omega(p; \nu)$ is defined as some weighting scheme that applies a one-parameter ethical weight ν (Donaldson and Weymark, 1980, 1983). It is defined accordingly as:

$$\omega(p; \nu) = \nu(1 - p)^{(\nu-1)} \quad \text{for } \nu \geq 1 \quad (4.35)$$

³⁹ Homotheticity of the social welfare functions is to ensure that they are consistent with relative inequality measures (Duclos *et al.*, 2003). It implies that the ratio of marginal social utilities of two individuals is constant when incomes are changed by the same proportions.

where ν is a parameter of aversion to “rank inequality” or reranking (Duclos *et al.*, 2003 p.73). Yaari (1988) interprets this as an indicator of the policy maker’s degree of equality mindedness. The larger its value, the higher the weights given to the inequalities of the poor and by implication, the faster is the fall in the weights $\omega(p; \nu)$ associated with higher ranks (p). Equation (4.34) represents the dual-parameter social welfare function (Duclos *et al.*, 2003) of pre-payment income. $X(p)$, ε and p remain as previously defined.

The utility function chosen (i.e. $U_\varepsilon(X)$) for $\varepsilon \geq 0$ is represented as:

$$U_\varepsilon(X) = \begin{cases} \frac{X^{1-\varepsilon}}{1-\varepsilon}, & \text{when } \varepsilon \neq 1 \\ \ln(X), & \text{when } \varepsilon = 1 \end{cases} \quad (4.36)$$

where $U_\varepsilon(X)$ is a concave isoelastic utility function. Concavity here implies that individuals are averse to uncertainty in their post-payment income levels (Duclos *et al.*, 2003) such that unequal treatment of equals would induce a loss in total welfare. The parameter ε can further be qualified as one of relative risk aversion (Atkinson, 1970).

Duclos *et al.* (2003) define $\xi_X(\varepsilon, \nu)$ as the *equally distributed equivalent (EDE) income* for a distribution of pre-payment income X , such that if everyone in the population gets $\xi_X(\varepsilon, \nu)$ it would generate the same level of social welfare as that generated by the current level of pre-payment income (X) distribution. Replacing $\xi_X(\varepsilon, \nu)$ in (4.34) they obtain

$$W_X(\varepsilon, \nu) = \int_0^1 U_\varepsilon(\xi_X(\varepsilon, \nu)) \omega(p; \nu) dp = U_\varepsilon(\xi_X(\varepsilon, \nu)). \quad (4.37)$$

The identity in (4.37) is obtained because $\xi_X(\varepsilon, \nu)$ is identical $\forall p$. Let us further denote the inverse utility function of (4.36) as:

$$U_\varepsilon^{-1}(X) = \begin{cases} (1-\varepsilon)X^{\frac{1}{1-\varepsilon}}, & \text{when } \varepsilon \neq 1 \\ \exp(X), & \text{when } \varepsilon = 1 \end{cases} \quad (4.38)$$

By inspection, we can rewrite (4.37) by inverse transformation as $\xi_X(\varepsilon, \nu) = U_\varepsilon^{-1}(W_X(\varepsilon, \nu))$.

Analogously, we can write the same expression for $\xi_N(\varepsilon, \nu)$, $\xi_N^E(\varepsilon, \nu)$, and $\xi_N^P(\varepsilon, \nu)$. Where $\xi_N(\varepsilon, \nu)$ is the *equally distributed equivalent income* for a distribution of post-payment income;

$\xi_N^E(\varepsilon, \nu)$ is the *equally distributed equivalent income* for a distribution of post-payment income where individuals at a rank p of pre-payment income are given their expected post payment income $\bar{N}(p)$ ⁴⁰ in formulating the social welfare function; and $\xi_N^P(\varepsilon, \nu)$ is the *equally distributed equivalent income* for a distribution of post-payment income where individuals at a rank p of pre-payment income are given the expected utility of their post-payment income $\overline{U_\varepsilon(N(p))}$ in formulating the appropriate social welfare function.

Following Atkinson (1970), inequality can be defined as the difference between $\xi_X(\varepsilon, \nu)$ and μ_X as a proportion of μ_X .

$$I_X(\varepsilon, \nu) = 1 - (\xi_X(\varepsilon, \nu) / \mu_X) \quad (4.39)$$

where $I_X(\varepsilon, \nu)$ is interpreted as the percentage of total income that could be spent in removing inequality and which would result in no loss in social welfare (Duclos *et al.*, 2003) and μ_X is the mean pre-payment income.

Similar inequality indices can be defined for $I_N(\varepsilon, \nu)$, $I_N^E(\varepsilon, \nu)$, and $I_N^P(\varepsilon, \nu)$. As noted earlier this class of social welfare functions is based on dual-parameters. They exhibit aversion to riskiness in post-payment incomes (through ε) and aversion to relative deprivation (through ν) (see Duclos *et al.*, 2003). Duclos *et al.* (2003) show that the following inequality $I_N(\varepsilon, \nu) \geq I_N^P(\varepsilon, \nu) \geq I_N^E(\varepsilon, \nu)$ holds. This can be inferred immediately from the construction of these indices and as a consequence of equation (4.39).

The redistributive effect of health care financing can then be decomposed as the difference between inequality in gross income and inequality in net income following the identity in (4.13) as:

$$\begin{aligned} \Delta I(\varepsilon, \nu) &= I_X(\varepsilon, \nu) - I_N(\varepsilon, \nu) \\ &= \underbrace{I_X(\varepsilon, \nu) - I_N^E(\varepsilon, \nu)}_{V^{DJA}} - \underbrace{(I_N^P(\varepsilon, \nu) - I_N^E(\varepsilon, \nu))}_{H^{DJA} \geq 0} - \underbrace{(I_N(\varepsilon, \nu) - I_N^P(\varepsilon, \nu))}_{R^{DJA} \geq 0} \end{aligned} \quad (4.40)$$

⁴⁰ Duclos *et al.* (2003) estimate this using a non-parametric (via Kernel function) approach.

where the constituent indices are as earlier defined. Equation (4.40) is referred to as the *DJA* decomposition of the redistributive effect of health care payments. Specifically V^{DJA} is the measure of vertical equity or progressivity (i.e. the decrease in inequality caused by a tax or payment that treats unequals unequally), H^{DJA} measures *classical* horizontal inequity (i.e. the increase in overall income inequality that is caused by unequal treatment of equals), and R^{DJA} measures reranking. The *DJA* decomposition framework has recently been used to analyse the redistributive effect of health care financing in Switzerland (Bilger, 2008) and Argentina (Cavagnero and Bilger, 2010).

When comparing the *DJA* with *AJL*, *VCL* and *UL*, it is important to note that the *DJA* decomposition is based on a fundamentally separate normative assumption (see Bilger, 2008). While *AJL*, *VCL* and *UL* require grouping of near-equals, *DJA* does not require this restriction but relies on the nonparametric estimation of the expected post-payment income function (Duclos *et al.*, 2003, Duclos and Araar, 2006). In fact the reranking component in *DJA* captures reranking more fully than in the *AJL* or *UL* formulation which groups near equals. However one of the difficulties with the use of the *DJA* framework is that it requires some normative choice of the values of ε and ν . Because ν is a parameter of aversion to reranking if $\nu = 1$, unitary weight is given to every household or individual irrespective of their rank. Therefore we would expect that $R^{DJA} = 0$. Increasing the value of ν is equivalent to weighing the reranking resentment of the poorest households or individuals. This may not necessarily increase R^{DJA} (Duclos *et al.*, 2003). However when $\nu = 2$ and $\varepsilon = 0$, $R^{DJA} \cong R^{AJL}$. This is so because as the bandwidth of close equals shrinks one should expect that $H^{AJL} \rightarrow 0$. When $\varepsilon = 0$ it means that there is no aversion to horizontal inequality and as such $H^{DJA} = 0$. As ε increases H^{DJA} also increases while V^{DJA} would increase if the tax or payment mechanism were progressive.

Empirically it has been shown that the value of ε should range between 0.25 and 1.0 while that of ν between 1 and 4 (Duclos, 2000, Duclos *et al.*, 2003). This conclusion is based on a “leaky bucket”/ efficiency loss experiment (see Duclos, 2000 who attributed this to Okun, 1975)⁴¹. In fact Duclos *et al.* (2003) suggest $\nu = 1.5$ and $\varepsilon = 0.4$ as ‘reasonable’ values for the

⁴¹ Basically the experiment is used to assess the range of values of ν and ε that are ethically ‘sensible’ (Duclos, 2000). It requires an assessment of society’s tolerance to the cost incurred in transferring income from a rich to a poor individual/household. This involves transferring an amount of money “from the rich to the poor, but a

parameters. All the studies that have used the DJA methodology in health care financing have set the parameters to these ‘reasonable’ values (see Table 4.10 for instance).

4.5 EMPIRICAL REVIEW OF REDISTRIBUTIVE EFFECT OF HEALTH CARE FINANCING

The earliest studies undertaken to analyse the redistributive effect of health care financing were conducted in developed economies. The first was in 1997 for the Netherlands (Wagstaff and van Doorslaer, 1997); the second in 1998 for Sweden (Gerdtham and Sundberg, 1998). Since then only a few more studies have been conducted and these have been concentrated within Europe (cf. van Doorslaer *et al.*, 1999, Bilger, 2008) or selected countries in Asia (cf. Wagstaff and van Doorslaer, 2001, Cavagnero and Bilger, 2010). The only study in Africa contained in Table 4.2 was conducted in Nigeria (Ichoku, 2006) but this only considered out-of-pocket payments and was restricted to only one state.

In empirical investigation of various studies, as summarised in Table 4.2, out-of-pocket payments generally induce a pro-rich redistribution, except for Argentina in 2002 where $RE > 0$. In Argentina the switch from a regressive relationship in 1997 to a progressive one by 2002 (a period of economic crisis in Argentina) was attributed to pro-poor government safety net programmes that shielded the poor from paying for health care. It is also likely that poor households forewent seeking care (Cavagnero and Bilger, 2010).

In almost all the cases presented in Table 4.2, the vertical effect dominates the horizontal and re-ranking effects. The result is consistent across the methodologies. This implies that the pro-rich redistribution induced by out-of-pocket payment is largely attributable to unequal (unfavourable) treatment of unequals. This can also be observed from the values of $V < 0$ indicating that out-of-pocket payments are generally regressive. In the case of Vietnam, Wagstaff and van Doorslaer (2001) estimate RE between 1993 and 1998 on both per capita pre-payment income and a measure of ability-to-pay that deducts household per capita food spending from household per capita income. Their results as shown in Table 4.2 indicate that

certain fraction of it is lost when doing so, for instance because of administrative costs. The extent of the loss, or *leakage*, in the transfer that society can accept determines the level of inequality aversion” (Pirttilä and Uusitalo, 2010 p.60).

assessing RE on the measure of ability-to-pay resulted in the same declining trend in RE as obtained using per capita income. Though the redistributive effect declined between the two periods, the contribution of V to RE increased in 1998 using the measure of ability-to-pay (Wagstaff and van Doorslaer, 2001) and was attributed to the increase in regressivity (V) of out-of-pocket payments during this period. In Nigeria using both the AJL approach and DJA approach, while out-of-pocket payments appear to be mildly progressive, there is a pro-rich redistribution because of the combined effects of the horizontal and re-ranking components. The progressivity of out-of-pocket payment is attributable to the poor not seeking care because they cannot afford it (Ichoku, 2006).

Direct taxes are generally progressive ($V > 0$) and therefore result in pro-poor redistribution ($RE > 0$). In all the countries reviewed (see Table 4.3), using various methodologies, direct taxes show positive redistributive effects that are attributable mainly to the vertical effect. The contribution of the vertical effect to the total redistribution (V/RE) is generally greater than or equal to 100%. With the exception of Argentina, all the studies on direct taxes were conducted in OECD countries. In Argentina using the DJA methodology, the decrease in RE between 1997 and 2002 was associated with a reduction in the progressivity of direct taxes (Cavagnero and Bilger, 2010) which in turn was attributed to the traditional Phillip's curve hypothesis. Inflation in Argentina increased over that period but unemployment decreased thereby increasing the number of tax payers, included those in the poorer quantiles. RE also decreased in the Netherlands between 1987 and 1992 (Wagstaff and van Doorslaer, 1997, van Doorslaer *et al.*, 1999) and in Switzerland between 1992 and 1998 (van Doorslaer *et al.*, 1999, Bilger, 2008). In both cases the reduction in RE is attributable to reductions in progressivity (via V) of direct taxes.

Unlike direct taxes, indirect taxes induce pro-rich redistribution ($RE < 0$). The negative redistribution is associated mainly with the vertical effect. In all the countries listed in Table 4.4 indirect taxes are regressive ($V < 0$). For Argentina redistribution associated with indirect taxes increased slightly, though not significantly, in absolute terms between 1997 and 2002. This was due to slight increases in horizontal equity (H). Also in Switzerland and the Netherlands there were only small changes in the redistributive effects of indirect taxes between the periods studied. The positive redistributive effects induced by direct taxes dominate the negative redistributive effects of indirect taxes. This resulted in a generally pro-poor

redistributive effect ($RE > 0$) of general taxes as shown in Table 4.5. Again the vertical effect generally drives the results across all countries.

Redistribution associated with private insurance cover varies across countries depending on the nature of the insurance cover. As observed before, private insurance is regressive in countries like the United States, Switzerland, France and Ireland (see Table 4.6). In these countries, as shown in Table 4.6, private insurance induces pro-rich redistribution ($RE < 0$). Pro-poor redistribution was however observed in countries where private insurance is progressive. These results were observed mainly because the contribution of the vertical effect to overall redistribution is greater than the combined contributions of the horizontal and re-ranking effects. The regressive patterns that contribute to pro-rich redistribution in private insurance in some countries are likely to be due to premiums not being related to ability-to-pay (van Doorslaer *et al.*, 1999). Generally for countries such as Switzerland and the United States that have private cover as the sole form of cover for a large spectrum of the population, the effect of this on redistribution is clear. Though Germany operates private insurance cover, this is purchased mainly by the rich and accounts for the positive value of V and an overall pro-poor redistribution. The case of Ireland is important to note. Private insurance is regressive and induces a negative redistribution (van Doorslaer *et al.*, 1999). During the period of the study Ireland operated a state-backed non-profit insurance - the Voluntary Health Insurance (VHI) - and by 1987 about 34% of the population had such cover (Harmon and Nolan, 2001). This mainly provided cover to the top earning Irish population. Within the top earning Irish population because the premiums are not income-related, the regressive pattern observed is a reflection of the distribution among the insured.

Social insurance also varies in terms of its redistributive power. Apart from Argentina where the combined effects of H and R outweigh V , redistribution is generally attributed to the vertical effect (i.e. progressivity) discussed earlier. Countries such as Switzerland, Germany and Netherlands have both regressive and pro-rich redistribution associated with social insurance. In Switzerland in 1992 van Doorslaer *et al.* (1999) estimate a pro-poor redistributive effect while in 1998 Bilger (2008) estimates a pro-rich redistribution. The switch from pro-poor to pro-rich is attributed to the fact that social health insurance was regressive in 1998. In fact the vertical inequity observed in 1998 was associated with differential premiums that are not related to income (Bilger, 2008). In the Netherlands Wagstaff and van Doorslaer (1997) estimate a

progressive pattern for the AWBZ scheme⁴² and a regressive pattern for the sickness funds⁴³ (Wagstaff and van Doorslaer, 1997). Though contributions to the AWBZ are a fixed proportion of income, the summary in Table 4.7 indicates that H is not zero. The non-zero value of H has been attributed to the operational and institutional features of the scheme. Because contributions are by individuals and the amounts do not take into account for example differences in household size (Wagstaff and van Doorslaer, 1997), differential treatments are likely to be observed in reality.

Redistribution associated with total health care financing, public health care financing and private health care financing varies depending on the relative importance of each mechanism and their relative vertical, horizontal and re-ranking effects. The combination of all private health care financing sources, on average, induced pro-rich redistribution ($RE < 0$) in all but one (the Netherlands) of the OECD countries (see Table 4.9). The pro-poor redistribution in the Netherlands arises because of the pro-poor redistribution in private insurance that dominates the pro-rich redistribution observed for out-of-pocket payments (van Doorslaer *et al.*, 1999). The redistributive effect associated with private health care financing is attributed largely to the unequal treatment of unequals – vertical effect. The redistributive effect associated with public financing sources is mixed across the OECD countries. While Germany, the Netherlands and Switzerland record pro-rich redistribution, the rest of the countries contained in Table 4.8 show pro-poor redistribution (van Doorslaer *et al.*, 1999) and these are again largely attributed to the vertical effect. In total, as summarised in Table 4.10, apart from a few countries, health care financing across many countries result in a pro-rich redistribution. In Argentina an increase in pro-poor redistribution was observed in 2002 (Cavagnero and Bilger, 2010). This resulted from a progressive general tax, private insurance and out-of-pocket payments. Indeed the last became much more progressive by 2002.

⁴² This is a social insurance scheme that provides cover to everyone for catastrophic and related costs. Only individuals below the pension age make contributions which are a flat percentage of earnings with a maximum ceiling (Wagstaff and van Doorslaer, 1997).

⁴³ This is another form of social insurance that provides cover to enrollees for short term medical care. Contributions are however proportional to earnings up to a specified ceiling. Enrolment is compulsory for individuals earning below a specified amount (Wagstaff and van Doorslaer, 1997).

4.6 GAPS IN THE LITERATURE

This review of the literature has shown that most of the studies assessing equity in health care financing are conducted in developed countries. This is more so for studies that assess the income redistributive effect of health care financing. In fact there is no study in Africa that comprehensively addresses these issues. It should be noted that issues of equity in health care financing should not only be restricted to developed countries alone.

On a methodological front, while there are competing methodologies for assessing the income redistributive effect associated with financing health care, there has not been a study that assess these methods in relation to which is applicable in situations of high income inequality. This is largely because the application of these methods has been based on developed economies that have relatively low income inequalities.

Again, it is well known that economies of scale exist in households. Within the context of Africa, the existence of what Ichoku *et al.* (2012) call ‘community spirit’ means that economies of scale may be relevant. However, there has not been any study that assesses the impact of changing assumptions about economies of scale, and more broadly, equivalence scale, on the measures of progressivity or redistribution associated with health care financing.

Domestically, there is a need for evidence based information for the current move towards a National Health Insurance in South Africa. Unfortunately, this review did not uncover any previous study on the relative progressivity and the redistributive impact of the past or current health care financing arrangements.

This study attempts to fill these gaps in what follows in subsequent chapters.

Table 4.1: Summary of Kakwani indices of progressivity across studies

Country	Author(s)	Period	Out-of-pocket payment	Indirect taxes	Direct taxes	General Taxes	Social insurance	Private insurance	Total Public finance	Total private finance	Other	Total Health care finance
Australia	Lairson et al. (1995)	1988/89	-0.150	-	-	0.07	-	0.09	-	-	-0.42 ^a	0.010
Bangladesh	O'Donnell et al. (2008a)	1999/2000	0.2192	0.111	0.5523	-	-	-	-	-	-	0.2142
Brazil ^p	Ugá and Santos (2007)	2002	Regressive	-	-	Progressive	-	Progressive	-	-	-	Proportional
China ^z	O'Donnell et al. (2008a)	2000	-0.0168	0.0398	0.1521	-	0.2348	-	-	-	-	0.0404
Colombia ^t	Castano et al. (2002)	1984	-0.1257	-	-	-	-	-	-	-	-	-
Colombia ^t	Castano et al. (2002)	1997	-0.3498	-	-	-	-	-	-	-	-	-
Colombia ^v	Castano et al. (2002)	1984	-0.0092	-	-	-	-	-	-	-	-	-
Colombia ^u	Castano et al. (2002)	1997	0.0026	-	-	-	-	-	-	-	-	-
Côte d'Ivoire ^c	Cissé et al. (2007)	1998/99	-0.15	-	-	-	-	-	-	-	-	-
Denmark	Wagstaff et al. (1999b)	1981	-0.2568	-0.1498	0.0431	-0.0022	-	0.0477	-0.0022	-0.2350	-	-0.0348
Denmark	Wagstaff et al. (1999b)	1987	-0.2654	-0.1126	0.0624	0.0372	-	0.0313	0.0372	-0.2363	-	-0.0047
Egypt	O'Donnell et al. (2008b)	1997	0.0644	0.1435	0.2501	-	-0.0532	-0.0011	-	-	-0.0061 ^b	0.0819
Finland	Wagstaff et al. (1999b)	1990	-0.2419	-0.0969	0.1272	0.0555	0.0937	0.0000	0.0604	-0.2419	-	0.0181
Finland ^e	Klavus (2001)	1987	Regressive	Regressive	-	-	-	-	-	-	Progressive ^f	Regressive
Finland ^e	Klavus (2001)	1996	Regressive	Regressive	-	-	-	-	-	-	Progressive ^f	Regressive
France	Wagstaff et al. (1999b)	1984	-0.2453	-	-	-	0.0960	-0.1788	0.0960	-0.2228	-	0.0236
France	Wagstaff et al. (1999b)	1989	-0.3396	-	-	-	0.1112	-0.1956	0.1112	-0.3054	-	0.0012
Germany	Wagstaff et al. (1999b)	1989	-0.0963	-0.0922	0.2488	0.1100	-0.0977	0.1219	-0.0533	-0.0067	-	-0.0452
Ghana	Akazili (2010); Akazili <i>et al.</i> (2011)	2005/06	-0.0680	0.057	0.202	Progressive ^y	-	-	-	-0.0680	0.144 ^g	0.071
Guinea	Cissé et al. (2007)	1998/99	-0.5200	-	-	-	-	-	-	-	-	-
Hong Kong SAR	O'Donnell et al. (2008a)	1999/2000	0.0113	0.1102	0.394	-	-	0.0403	-	-	-	0.1689
India ⁿ	O'Donnell et al. (2008a)	1999/2000	0.0461	0.0579	-	-	-	-	-	-	-	0.0485
Indonesia	O'Donnell et al. (2008a)	2001	0.1761	0.0741	0.1962	-	0.3057	-	-	-	-	0.1729
Ireland	Smith (2010)	1987/88	-0.055	-0.13	-	0.05	0.175	0.08	-	-	-	0.050
Ireland	Smith (2010)	1999/2000	-0.100	-0.13	-	0.04	0.18	0.058	-	-	-	0.040
Ireland	Smith (2010)	2004/05	-0.108	-	-	-	0.15	-0.032	-	-	-	-
Ireland	Wagstaff et al. (1999b)	1987	-0.1472	-	0.2666	-	0.1263	-0.0210	-	-0.0965	-	-
Italy	Wagstaff et al. (1999b)	1991	-0.0807	-0.1135	0.1554	0.0343	0.1072	0.1705	0.0712	-0.0612	-	0.0413
Japan	O'Donnell et al. (2008a)	1998	-0.2691	-0.2232	0.095	-	-0.0415	-	-	-	-	-0.0688
Korea Rep. ^k	O'Donnell et al. (2008a)	2000	0.0124	0.0379	0.2683	-	-0.1634	-	-	-	-	-0.0239

Country	Author(s)	Period	Out-of-pocket payment	Indirect taxes	Direct taxes	General Taxes	Social insurance	Private insurance	Total Public finance	Total private finance	Other	Total Health care finance
Kyrgyz Rep.	O'Donnell et al. (2008a)	2000	-0.052	0.0508	0.2395	-	0.1422	-	-	-	-	0.0087
Malaysia	Yu et al., (2008, 2011)	1998/99	0.001	-0.0179	0.1474	-	0.0004	0.0857	-	-	-	0.2166
Mali ^m	Cissé et al. (2007)	1998/99	-0.220	-	-	-	-	-	-	-	-	-
Nepal	O'Donnell et al. (2008a)	1995/96	0.0533	0.1143	0.1436	-	-	-	-	-	-	0.0625
Netherlands	Wagstaff et al. (1999b)	1987	-0.0235	-0.1109	0.2428	0.0712	-0.1632	0.0859	-0.1309	0.0472	-	-0.0835
Netherlands	Wagstaff et al. (1999b)	1992	-0.0377	-0.0885	0.2003	0.0714	-0.1286	0.0833	-0.1003	0.0434	-	-0.0703
Palestine ^p	Abu-Zaineh et al. (2008)	2004	-0.0830	-	-	-	0.0603	0.1095	-	-	-	-0.0677
Palestine ^q	Abu-Zaineh et al. (2008)	2004	-0.0636	-	-	-	0.0524	0.1666	-	-	-	-0.0473
Philippines	O'Donnell et al. (2008a)	1999	0.1391	0.0024	0.3809	-	0.2048	0.1199	-	-	-	0.1631
Portugal	Wagstaff et al. (1999b)	1990	-0.2424	-0.0347	0.2180	0.0601	0.1845	0.1371	0.0723	-0.2287	-	-0.0445
Senegal ^s	Cissé et al. (2007)	1998/99	-0.080	-	-	-	-	-	-	-	-	-
Spain	Wagstaff et al. (1999b)	1980	0.0079	0.0107	0.1121	0.0706	-0.1117	-0.0537	-0.0780	0.0003	-	-0.0589
Spain	Wagstaff et al. (1999b)	1990	-0.1801	-0.1533	0.2125	0.0486	0.0615	-0.0224	0.0509	-0.1627	-	0.0004
Sri Lanka ^L	O'Donnell et al. (2008a)	1996/97	0.0687	-0.010	0.5693	-	-	-	-	-	-	0.085
Sweden	Wagstaff et al. (1999b)	1980	-0.4033	-0.0888	0.0619	0.0402	0.0046	-	0.0325	-0.4033	-	-0.0028
Sweden	Wagstaff et al. (1999b)	1990	-0.2402	-0.0827	0.0529	0.0371	0.0100	-	0.0100	-0.2402	-	-0.0158
Switzerland	Wagstaff et al. (1999b)	1982	-0.3185	-0.0652	0.1958	0.1463	0.0145	-0.2443	0.1414	-0.2675	-	-0.1019
Switzerland	Wagstaff et al. (1999b)	1992	-0.3619	-0.0722	0.2055	0.1590	0.0551	-0.2548	0.1389	-0.2945	-	-0.1402
Taiwan	O'Donnell et al. (2008a)	2000	-0.0962	0.0296	0.2601	-	-0.0305	0.1961	-	-	-	-0.0119
Tanzania	Mtei and Borghi (2010)	2005	-0.08	0.07	0.48	0.18	-	-0.49	-	-	-	0.05
Thailand	O'Donnell et al. (2008a)	2002	0.0907	0.1819	0.5101	-	0.1803	0.0039	-	-	-	0.1972
UK	Wagstaff et al. (1999b)	1993	-0.2229	-0.1522	0.2843	0.0456	0.1867	0.0766	0.0792	-0.0919	-	0.0510
US	Wagstaff et al. (1999b)	1987	-0.3874	-0.0674	0.2104	0.1487	0.0181	-0.2374	0.1060	-0.3168	-	-0.1303
Vietnam	Wagstaff & van Doorslaer (2001)	1993	-0.0537	-	-	-	-	-	-	-	-	-
Vietnam	Wagstaff & van Doorslaer (2001)	1998	0.0053	-	-	-	-	-	-	-	-	-
Zambia	Mwenge (2010)	1998	0.03656	-	-	-	-	-	-	-	-	-
Zambia	Mwenge (2010)	2004	-0.0799	-	-	-	-	-	-	-	-	-
Zambia	Mwenge (2010)	2006	0.0171	-	-	-	-	-	-	-	-	-

^aWorkers compensation, vehicle 3rd party; ^bEarmarked tax (cigarette tax); ^conly the capital (Abidjan); ^dpayment shares across deciles of per capita income was used to assess progressivity;

^eProgressive dominance was used to assess progressivity; ^fSickness fund/municipality local tax; ^gNational health insurance contributions; ^honly urban population; ⁱthis excludes the warring northern province; ⁿonly the capital (Bamako); ^sonly Punjab state; ^ponly West Bank; ^qonly Gaza Strip; ^sonly Dakar; ^rusing income; ^uusing household expenditures; ^wonly the capital (Conakry);

^vpayment shares across quintiles of per capita consumption expenditure was used; ^zonly 10 provinces were covered.

Table 4.2: Redistributive effect associated with out-of-pocket payment

Country	Author(s)	Period	Method [†]	V	(V/RE)%	H	(H/RE)%	R	(R/RE)%	RE
Argentina ^{A, B}	Cavagnero and Bilger (2010)	1997	DJA	0.00261	-82.3%	0.00327	-103.2%	0.00250	-78.9%	-0.00317
Argentina ^{A, B}	Cavagnero and Bilger (2010)	2002	DJA	0.01508	162.0%	0.00362	38.9%	0.00215	23.1%	0.00931
Denmark	van Doorslaer et al. (1999)	1987	AJL	-0.0021	100.0%	-	-	-	-	-0.0021
Finland	van Doorslaer et al. (1999)	1990	AJL	-0.0037	94.9%	-	-	-	-	-0.0039
France	van Doorslaer et al. (1999)	1989	AJL	-0.0031	93.9%	0.0004	-28.6%	-	-	-0.0033
Germany	van Doorslaer et al. (1999)	1988	AJL	-0.001	71.4%	0.0001 ^x	-6.7% ^y	-	-	-0.0014
Ireland	van Doorslaer et al. (1999)	1987	AJL	-0.0014	93.3%	0.0006 ^x	-26.1% ^y	-	-	-0.0015
Italy	van Doorslaer et al. (1999)	1991	AJL	-0.0017	73.9%	0.0001 ^x	-20.0% ^y	-	-	-0.0023
Netherlands	van Doorslaer et al. (1999)	1992	AJL	-0.0004	80.0%	0.0004 ^x	-28.6% ^y	-	-	-0.0005
Netherlands	Wagstaff and van Doorslaer (1997)	1987	AJL	-0.00081	81.8%	0.00011	-11.1%	0.00027	-27.3%	-0.00099
Nigeria ^N	Ichoku, (2006)	2004	AJL	0.0136	-1700.0%	0.007	-875.0%	0.0074	-925.0%	-0.0008
Nigeria ^{N, B}	Ichoku, (2006)	2004	DJA	0.0284	-1352.4%	0.0202	-961.9%	0.0103	-490.5%	-0.0021
*Palestine ^G	Abu-Zaineh et al. (2009)	2004	UL	-0.00840	33.9%	0.00061	-2.5%	0.01576	-63.6%	-0.02477
Palestine	Abu-Zaineh et al. (2008)	2004	RS	-	-	-	-	-	-	-0.0247
**Palestine ^W	Abu-Zaineh et al. (2009)	2004	UL	-0.01580	42.7%	0.0006	-1.6%	0.02060	-55.7%	-0.0370
Palestine ^W	Abu-Zaineh et al. (2008)	2004	RS	-	-	-	-	-	-	-0.0370
Portugal	van Doorslaer et al. (1999)	1990	AJL	-0.0084	75.7%	-	-	-	-	-0.0111
Sweden	van Doorslaer et al. (1999)	1990	AJL	-0.0004	80.0%	-	-	-	-	-0.0005
Switzerland	van Doorslaer et al. (1999)	1992	AJL	-0.0011	100.0%	0.0000	0.0%	0.0000	0.0%	-0.0011
Switzerland ^B	Bilger (2008)	1998	DJA	-0.00427	77.8%	0.00068	-12.4%	0.00054	-9.8%	-0.00549
United Kingdom	van Doorslaer et al. (1999)	1992	AJL	-0.0017	81.0%	-	-	-	-	-0.0021
United States	van Doorslaer et al. (1999)	1987	AJL	-0.0109	85.2%	0.0019 ^x	-14.8% ^y	-	-	-0.0128
Vietnam	Wagstaff & van Doorslaer (2001)	1993	AJL	-0.0027	47.4%	0.001	-24.6%	0.0016	-28.1%	-0.0057
Vietnam	Wagstaff & van Doorslaer (2001)	1998	AJL	-0.0001	4.2%	0.001	-41.7%	0.0013	-54.2%	-0.0024
Vietnam ^C	Wagstaff & van Doorslaer (2001)	1993	AJL	-0.0116	41.9%	0.009	-31.8%	0.0073	-26.4%	-0.0277
Vietnam ^C	Wagstaff & van Doorslaer (2001)	1998	AJL	-0.0109	62.6%	0.003	-19.0%	0.0032	-18.4%	-0.0174

[†] AJL = Aronson-Johnson-Lambert methodology, DJA = Duclos-Jalbert-Araar methodology, RS = Reynolds-Smolensky methodology, UL = Urban-Lambert methodology; *within group, between group and entire group reranking indices are 0.003002, 0.01255 and 0.00021 respectively; **within group, between group and entire group reranking indices are 0.002402, 0.01793 and 0.00027 respectively; ^Aonly urban households; ^Bthe values of DJA parameters are: $\varepsilon = 0.4$, $\nu = 1.5$; here pre-payment income was replaced with a measure of ability-to-pay derived as pre-payment income less household actual food outlay; ^Conly Gaza Strip; ^Nonly one state in Nigeria; ^Wonly West Bank; ^xrepresents (H+R); ^yrepresents ((H+R)/RE)%.

Table 4.3: Redistributive effect associated with direct taxes

Country	Author(s)	Period	Method [†]	V	(V/RE)%	H+R	((H+R)/RE)%	RE
Argentina ^{A,B}	Cavagnero and Bilger (2010)	1997	DJA	0.00383	143.4%	0.00116	43.4	0.00267
Argentina ^{A,B}	Cavagnero and Bilger (2010)	2002	DJA	0.00223	119.3%	0.00036	19.3%	0.00187
Denmark	van Doorslaer <i>et al.</i> (1999)	1987	AJL	0.0028	103.7%	0.0001	3.7%	0.0027
Finland	van Doorslaer <i>et al.</i> (1999)	1990	AJL	0.0070	101.4%	0.0001	1.4%	0.0069
Germany	van Doorslaer <i>et al.</i> (1999)	1988	AJL	0.0024	100.0%	0.0000	0.0%	0.0024
Ireland	van Doorslaer <i>et al.</i> (1999)	1987	AJL	0.0056	100.0%	0.0000	0.0%	0.0056
Italy	van Doorslaer <i>et al.</i> (1999)	1991	AJL	0.0025	100.0%	0.0000	0.0%	0.0025
Netherlands	Wagstaff and van Doorslaer (1997)	1987	AJL	0.00242	100.4%	-	-	0.00241
Netherlands	van Doorslaer <i>et al.</i> (1999)	1992	AJL	0.0014	107.7%	0.0001	7.7%	0.0013
Portugal	van Doorslaer <i>et al.</i> (1999)	1990	AJL	0.0038	111.8%	0.0004	11.8%	0.0034
Sweden	van Doorslaer <i>et al.</i> (1999)	1990	AJL	0.0054	103.8%	0.0002	3.8%	0.0052
Switzerland	van Doorslaer <i>et al.</i> (1999)	1992	AJL	0.0041	102.5%	0.0001	2.5%	0.0040
Switzerland ^B	Bilger (2008)	1998	DJA	0.00152	111.8%	0.00016	11.8%	0.00136
United Kingdom	van Doorslaer <i>et al.</i> (1999)	1992	AJL	0.0043	102.4%	0.0001	2.4%	0.0042
United States	van Doorslaer <i>et al.</i> (1999)	1987	AJL	0.0065	103.2%	0.0002	3.2%	0.0063

[†] AJL = Aronson-Johnson-Lambert methodology, DJA = Duclos-Jalbert-Araar methodology; ^Aonly urban households; ^Bthe values of DJA parameters are: $\varepsilon = 0.4$, $\nu = 1.5$.

Table 4.4: Redistributive effect associated with indirect taxes

Country	Author(s)	Period	Method [†]	V	(V/RE)%	H+R	((H+R)/RE)%	RE
Argentina ^{^,b}	Cavagnero and Bilger (2010)	1997	DJA	-0.00066	71.7%	0.00026	-28.3%	-0.00092
Argentina ^{^,b}	Cavagnero and Bilger (2010)	2002	DJA	-0.00067	60.4%	0.00043	-38.7%	-0.00111
Denmark	van Doorslaer <i>et al.</i> (1999)	1987	AJL	-0.0003	100.0%	0.0000	0.0%	-0.0003
Finland	van Doorslaer <i>et al.</i> (1999)	1990	AJL	-0.0024	96.0%	0.0001	-4.0%	-0.0025
Germany	van Doorslaer <i>et al.</i> (1999)	1988	AJL	-0.0007	100.0%	0.0000	0.0%	-0.0007
Italy	van Doorslaer <i>et al.</i> (1999)	1991	AJL	-0.0014	100.0%	0.0000	0.0%	-0.0014
Netherlands	van Doorslaer <i>et al.</i> (1999)	1992	AJL	-0.0003	100.0%	0.0000	0.0%	-0.0003
Netherlands	Wagstaff and van Doorslaer (1997)	1987	AJL	-0.0006	100.0%	0.0000	0.0%	-0.0006
Portugal	van Doorslaer <i>et al.</i> (1999)	1990	AJL	-0.0011	91.7%	0.0001	-8.3%	-0.0012
Sweden	van Doorslaer <i>et al.</i> (1999)	1990	AJL	-0.0011	100.0%	0.0000	0.0%	-0.0011
Switzerland	van Doorslaer <i>et al.</i> (1999)	1992	AJL	-0.0005	100.0%	0.0000	0.0%	-0.0005
Switzerland ^b	Bilger (2008)	1998	DJA	-0.00042	85.7%	0.00007	-14.3%	-0.00049
United Kingdom	van Doorslaer <i>et al.</i> (1999)	1992	AJL	-0.0017	94.4%	0.0001	-5.6%	-0.0018
United States	van Doorslaer <i>et al.</i> (1999)	1987	AJL	-0.0006	85.7%	0.0001	-14.3%	-0.0007

[†] AJL = Aronson-Johnson-Lambert methodology, DJA = Duclos-Jalbert-Araar methodology; [^]only urban households; ^bthe values of DJA parameters are: $\varepsilon = 0.4$, $\nu = 1.5$.

Table 4.5: Redistributive effect associated with general taxes

Country	Author(s)	Period	Method [†]	V	(V/RE)%	H+R	((H+R)/RE)%	RE
Denmark	van Doorslaer <i>et al.</i> (1999)	1987	AJL	0.0025	104.2%	0.0001	4.2%	0.0024
Finland	van Doorslaer <i>et al.</i> (1999)	1990	AJL	0.0046	104.5%	0.0002	4.5%	0.0044
Germany	van Doorslaer <i>et al.</i> (1999)	1988	AJL	0.0018	100.0%	0.0000	0.0%	0.0018
Italy	van Doorslaer <i>et al.</i> (1999)	1991	AJL	0.0012	100.0%	0.0000	0.0%	0.0012
Netherlands	van Doorslaer <i>et al.</i> (1999)	1992	AJL	0.0010	100.0%	0.0000	0.0%	0.0010
Portugal	van Doorslaer <i>et al.</i> (1999)	1990	AJL	0.0028	93.3%	-0.0002	-6.7%	0.0030
Switzerland	van Doorslaer <i>et al.</i> (1999)	1992	AJL	0.0036	102.9%	0.0001	2.9%	0.0035
Switzerland [‡]	Bilger (2008)	1998	DJA	0.00103	118.4%	0.00017	19.0%	0.00087
United Kingdom	van Doorslaer <i>et al.</i> (1999)	1992	AJL	0.0045	102.3%	0.0001	2.3%	0.0044
United States	van Doorslaer <i>et al.</i> (1999)	1987	AJL	0.0062	105.1%	0.0003	5.1%	0.0059

[†]AJL = Aronson-Johnson-Lambert methodology, DJA = Duclos-Jalbert-Araar methodology; [‡]the values of DJA parameters are: $\varepsilon = 0.4$, $\nu = 1.5$.

Table 4.6: Redistributive effect associated with private health insurance

Country	Author(s)	Period	Method [†]	V	(V/RE)%	H	(H/RE)%	R	(R/RE)%	RE
Argentina ^{A, B}	Cavagnero and Bilger (2010)	1997	DJA	0.00583	161.5%	0.00159	44.0%	0.00064	17.7%	0.00361
Argentina ^{A, B}	Cavagnero and Bilger (2010)	2002	DJA	0.00836	155.4%	0.00246	45.7%	0.00052	9.7%	0.00538
Denmark	van Doorslaer <i>et al.</i> (1999)	1987	AJL	0.0000	0.0%	0.0000 ^x	0.0% ^y	-	-	0.0000
France	van Doorslaer <i>et al.</i> (1999)	1989	AJL	-0.0027	93.1%	0.0002 ^x	-6.9% ^y	-	-	-0.0029
Germany	van Doorslaer <i>et al.</i> (1999)	1988	AJL	0.0013	130.0%	0.0003 ^x	30.0% ^y	-	-	0.001
Ireland	van Doorslaer <i>et al.</i> (1999)	1987	AJL	-0.0001	100.0%	0.0000 ^x	0.0% ^y	-	-	-0.0001
Italy	van Doorslaer <i>et al.</i> (1999)	1991	AJL	0.0002	100.0%	0.0000 ^x	0.0% ^y	-	-	0.0002
Netherlands	van Doorslaer <i>et al.</i> (1999)	1992	AJL	0.0014	127.3%	0.0003 ^x	27.3% ^y	-	-	0.0011
Netherlands	Wagstaff and van Doorslaer (1997)	1987	AJL	0.00133	149.4%	0.00019	21.3%	0.00045	50.6%	0.00089
*Palestine ^G	Abu-Zaineh <i>et al.</i> (2009)	2004	UL	0.00014	113.6%	0.000004	3.2%	0.000013	10.4%	0.000125
Palestine ^G	Abu-Zaineh <i>et al.</i> (2008)	2004	RS	-	-	-	-	-	-	0.0001
**Palestine ^w	Abu-Zaineh <i>et al.</i> (2009)	2004	UL	0.00018	128.6%	0.000002	1.4%	0.000038	27.1%	0.00014
Palestine ^w	Abu-Zaineh <i>et al.</i> (2008)	2004	RS	-	-	-	-	-	-	0.0001
Portugal	van Doorslaer <i>et al.</i> (1999)	1990	AJL	0.0002	200.0%	0.0001 ^x	100.0% ^y	-	-	0.0001
Switzerland	van Doorslaer <i>et al.</i> (1999)	1992	AJL	-0.0121	97.6%	0.0003 ^x	-2.4% ^y	-	-	-0.0124
Switzerland ^B	Bilger (2008)	1998	DJA	-0.00129	83.2%	0.00017	-11.0%	0.009	-5.8%	-0.00155
United Kingdom	van Doorslaer <i>et al.</i> (1999)	1992	AJL	0.0002	200.0%	0.0001 ^x	100.0% ^y	-	-	0.0001
United States	van Doorslaer <i>et al.</i> (1999)	1987	AJL	-0.0081	92.0%	0.0007 ^x	-8.0% ^y	-	-	-0.0088

[†]AJL = Aronson-Johnson-Lambert methodology, DJA = Duclos-Jalbert-Araar methodology, RS = Reynolds-Smolensky methodology, UL = Urban-Lambert methodology; ^A only urban households; ^B the values of DJA parameters are: $\varepsilon = 0.4$, $\nu = 1.5$; ^G only Gaza Strip; ^w only West Bank; ^x represents (H+R); ^y represents ((H+R)/RE)%; * within group, between group and entire group reranking indices are 0.00001, 0.000003 and 0.000 respectively; ** within group, between group and entire group reranking indices are 0.000022, 0.000016 and 0.000 respectively.

Table 4.7: Redistributive effect associated with social health insurance

Country	Author(s)	Period	Method [†]	V	(V/RE)%	H	(H/RE)%	R	(R/RE)%	RE
Argentina ^{A, B}	Cavagnero and Bilger (2010)	1997	DJA	0.0005	-454.5%	0.00009	-81.8%	0.00051	-463.6%	-0.00011
Argentina ^{A, B}	Cavagnero and Bilger (2010)	2002	DJA	0.00221	-334.8%	0.00206	-312.1%	0.00081	-122.7%	-0.00066
Finland	van Doorslaer <i>et al.</i> (1999)	1990	AJL	0.0011	100.0%	0.0000 ^x	0.0% ^y	-	-	0.0011
France	van Doorslaer <i>et al.</i> (1999)	1989	AJL	0.0165	120.4%	0.0028 ^x	20.4% ^y	-	-	0.0137
Germany	van Doorslaer <i>et al.</i> (1999)	1988	AJL	-0.0072	86.7%	0.0011 ^x	-13.3% ^y	-	-	-0.0083
Ireland	van Doorslaer <i>et al.</i> (1999)	1987	AJL	0.0006	100.0%	0.0000 ^x	0.0% ^y	-	-	0.0006
Italy	van Doorslaer <i>et al.</i> (1999)	1991	AJL	0.0061	107.0%	0.0004 ^x	7.0% ^y	-	-	0.0057
Netherlands	van Doorslaer <i>et al.</i> (1999)	1992	AJL	-0.01	95.2%	0.0005 ^x	-4.8% ^y	-	-	-0.0105
Netherland [*]	Wagstaff and van Doorslaer (1997)	1987	AJL	-0.01164	90.4%	0.00029	-2.3%	0.00094	-7.3%	-0.01288
Netherlands ^{**}	Wagstaff and van Doorslaer (1997)	1987	AJL	0.00133	111.8%	0.00008	6.7%	0.00007	5.9%	0.00119
Portugal	van Doorslaer <i>et al.</i> (1999)	1990	AJL	0.0010	142.9%	0.0003 ^x	42.9% ^y	-	-	0.0007
Sweden	van Doorslaer <i>et al.</i> (1999)	1990	AJL	0.0003	100.0%	0.0000 ^x	0.0% ^y	-	-	0.0003
Switzerland	van Doorslaer <i>et al.</i> (1999)	1992	AJL	0.0001	100.0%	0.0000 ^x	0.0% ^y	-	-	0.0001
Switzerland ^b	Bilger (2008)	1998	DJA	-0.00987	98.6%	0.00002	-0.2%	0.00011	-1.1%	-0.01001
United Kingdom	van Doorslaer <i>et al.</i> (1999)	1992	AJL	0.0018	105.9%	0.0001 ^x	5.9% ^y	-	-	0.0017
United States	van Doorslaer <i>et al.</i> (1999)	1987	AJL	0.0003	150.0%	0.0001 ^x	50.0% ^y	-	-	0.0002

[†]AJL = Aronson-Johnson-Lambert methodology, DJA = Duclos-Jalbert-Araar methodology; ^Aonly urban households; ^Bthe values of DJA parameters are: $\varepsilon = 0.4$, $\nu = 1.5$; ^xrepresents (H+R); ^yrepresents ((H+R)/RE)%; ^{*}Sickness fund (a form of social insurance); provides cover for only short term medical care, ^{**}AWBZ (a form of social insurance) that provides cover to everyone for catastrophic and related costs.

Table 4.8: Redistributive effect associated with total public finance

Country	Author(s)	Period	Method [†]	V	(V/RE)%	H+R	((H+R)/RE)%	RE
Denmark	van Doorslaer <i>et al.</i> (1999)	1987	AJL	0.0025	104.2%	0.0001	4.2%	0.0024
Finland	van Doorslaer <i>et al.</i> (1999)	1990	AJL	0.0058	105.5%	0.0003	5.5%	0.0055
Germany	van Doorslaer <i>et al.</i> (1999)	1988	AJL	-0.0054	62.8%	0.0032	-37.2%	-0.0086
Italy	van Doorslaer <i>et al.</i> (1999)	1991	AJL	0.0076	108.6%	0.0006	8.6%	0.0070
Netherlands	van Doorslaer <i>et al.</i> (1999)	1992	AJL	-0.0089	92.7%	0.0007	-7.3%	-0.0096
Portugal	van Doorslaer <i>et al.</i> (1999)	1990	AJL	0.0039	130.0%	0.0009	30.0%	0.0030
Switzerland	van Doorslaer <i>et al.</i> (1999)	1992	AJL	0.0036	100.0%	0.0000	0.0%	0.0036
Switzerland ^b	Bilger (2008)	1998	DJA	-0.00896	95.7%	0.00040	-4.3%	-0.00936
United Kingdom	van Doorslaer <i>et al.</i> (1999)	1992	AJL	0.0045	102.3%	0.0001	2.3%	0.0044
United States	van Doorslaer <i>et al.</i> (1999)	1987	AJL	0.0066	104.8%	0.0003	4.8%	0.0063

[†]AJL = Aronson-Johnson-Lambert methodology, DJA = Duclos-Jalbert-Araar methodology; ^bthe values of DJA parameters are: $\varepsilon = 0.4$, $\nu = 1.5$

Table 4.9: Redistributive effect associated with total private finance

Country	Author(s)	Period	Method [†]	V	(V/RE)%	H+R	((H+R)/RE)%	RE
Denmark	van Doorslaer <i>et al.</i> (1999)	1987	AJL	-0.002	95.2%	-	-	-0.0021
Finland	van Doorslaer <i>et al.</i> (1999)	1990	AJL	-0.0037	94.9%	-	-	-0.0039
Germany	van Doorslaer <i>et al.</i> (1999)	1988	AJL	0.0003	-75.0%	0.0007	-175.0%	-0.0004
Italy	van Doorslaer <i>et al.</i> (1999)	1991	AJL	-0.0015	71.4%	0.0006	-28.6%	-0.0021
Netherlands	van Doorslaer <i>et al.</i> (1999)	1992	AJL	0.0011	220.0%	0.0006	120.0%	0.0005
Portugal	van Doorslaer <i>et al.</i> (1999)	1990	AJL	-0.0083	75.5%	-	-	-0.0110
Switzerland	van Doorslaer <i>et al.</i> (1999)	1992	AJL	-0.0133	97.8%	0.0003	-2.2%	-0.0136
Switzerland ^b	Bilger (2008)	1998	DJA	-0.00588	78.6%	0.00161	-21.5%	-0.00748
United Kingdom	van Doorslaer <i>et al.</i> (1999)	1992	AJL	-0.0016	80.0%	-	-	-0.0020
United States	van Doorslaer <i>et al.</i> (1999)	1987	AJL	-0.0196	87.9%	0.0027	-12.1%	-0.0223

[†]AJL = Aronson-Johnson-Lambert methodology, DJA = Duclos-Jalbert-Araar methodology; ^bthe values of DJA parameters are: $\varepsilon = 0.4$, $\nu = 1.5$;

Table 4.10: Redistributive effect associated with total health care financing

Country	Author(s)	Period	Method [†]	V	(V/RE)%	H	(H/RE)%	R	(R/RE)%	RE
Argentina ^{A, B}	Cavagnero and Bilger (2010)	1997	DJA	0.00935	298.7%	0.00239	76.4%	0.00383	122.4%	0.00313
Argentina ^{A, B}	Cavagnero and Bilger (2010)	2002	DJA	0.02696	149.8%	0.00573	31.8%	0.00323	17.9%	0.0180
Denmark	van Doorslaer <i>et al.</i> (1999)	1987	AJL	-0.0003	60.0%	-	-	-	-	-0.0005
Finland	van Doorslaer <i>et al.</i> (1999)	1990	AJL	0.0018	-	-	-	-	-	-
France	van Doorslaer <i>et al.</i> (1999)	1989	AJL	0.0102	145.7%	-	-	-	-	0.007
Germany	van Doorslaer <i>et al.</i> (1999)	1988	AJL	-0.0052	83.9%	0.0010 ^x	-16.1% ^y	-	-	-0.0062
Ireland	van Doorslaer <i>et al.</i> (1999)	1987	AJL	0.0049	-	-	-	-	-	-
Italy	van Doorslaer <i>et al.</i> (1999)	1991	AJL	0.0061	127.1%	0.0013 ^x	27.1% ^y	-	-	0.0048
Netherlands	van Doorslaer <i>et al.</i> (1999)	1992	AJL	-0.008	93.0%	0.0006 ^x	-7.0% ^y	-	-	-0.0086
Netherlands	Wagstaff and van Doorslaer (1997)	1987	AJL	-0.00838	86.3%	0.00029	-3.0%	0.00103	-10.6%	-0.00971
*Palestine ^G	Abu-Zaineh <i>et al.</i> (2009)	2004	UL	-0.00730	28.7%	0.000703	-2.8%	0.017401	-68.5%	-0.02541
Palestine ^G	Abu-Zaineh <i>et al.</i> (2008)	2004	RS	-	-	-	-	-	-	-0.0254
**Palestine ^w	Abu-Zaineh <i>et al.</i> (2009)	2004	UL	-0.01470	38.8%	0.000602	-1.6%	0.022605	-59.6%	-0.03791
Palestine ^w	Abu-Zaineh <i>et al.</i> (2008)	2004	RS	-	-	-	-	-	-	-0.0379
Portugal	van Doorslaer <i>et al.</i> (1999)	1990	AJL	-0.0048	53.3%	-	-	-	-	-0.009
Sweden	van Doorslaer <i>et al.</i> (1999)	1990	AJL	0.0042	105.0%	-	-	-	-	0.0040
Switzerland	van Doorslaer <i>et al.</i> (1999)	1992	AJL	-0.0099	97.1%	0.0003 ^x	-2.9% ^y	-	-	-0.0102
Switzerland ^B	Bilger (2008)	1998	DJA	-0.01571	-	0.0014	-7.6%	0.00125	-6.8%	-0.01835
United Kingdom	van Doorslaer <i>et al.</i> (1999)	1992	AJL	0.0029	-	-	-	-	-	-
United States	van Doorslaer <i>et al.</i> (1999)	1987	AJL	-0.0139	80.3%	0.0034 ^x	-19.7% ^y	-	-	-0.0173

[†] AJL = Aronson-Johnson-Lambert methodology, DJA = Duclos-Jalbert-Araar methodology, RS = Reynolds-Smolensky methodology, UL = Urban-Lambert methodology; ^A only urban households; ^B the values of DJA parameters are: $\varepsilon = 0.4$, $\nu = 1.5$; ^G only Gaza Strip; ^w only West Bank; ^x represents (H+R); ^y represents ((H+R)/RE)%; * within group, between group and entire group reranking indices are 0.00310, 0.0141 and 0.000201 respectively; ** within group, between group and entire group reranking indices are 0.002603, 0.019802 and 0.00020 respectively

Chapter 5

DATA AND METHODOLOGY OF DATA EXTRACTION

5.1 INTRODUCTION

This chapter provides a detailed overview of the data used for analysis including how data on the various components of health care financing were extracted. Details of all the assumptions used are also provided. The precise methodologies used in assessing equity in health care financing are however contained in the chapters that present the results.

5.2 DATA

The main data for this study is the 2005/2006 round of the South African Income and Expenditure Survey (IES) conducted by Statistics South Africa. The IES is a nationally representative cross-sectional survey conducted every five years in South Africa. The 2005/06 IES survey was conducted across all nine South African provinces between September 2005 and August 2006. The survey collects detailed information on items and services acquired by South African households including the various sources of income (including in-kind) and household expenditure patterns. Detailed item-by-item expenditure by participating households including all acquired goods and services for households' own consumption were collected (Statistics South Africa, 2008d). A total of 3,000 primary sampling units (PSUs) were selected based on the Statistics South Africa's master sample of the 2001 population census enumeration areas. Each month a random sample of 250 PSUs were selected. In total eight dwelling units were selected within each PSU bringing to 24,000 the total number of selected dwelling units over the 12 months period. However the final dataset contains information on about 21,100 dwelling units (i.e. households)⁴⁴ translating into a response rate of >87%. Design weights for the data have been adjusted to account for the non-response (Statistics South Africa, 2008b). Household weights (W_{HH}) were generated as: $W_{HH} = 1 / (P_{PSU} \cdot P_{HH} \cdot r_{HH})$ where P_{PSU} is

⁴⁴ This was due to errors in listing, vacant and unoccupied dwelling, etc.

the PSU inclusion probability, P_{HH} is the household inclusion probability per PSU, and r_{HH} is the response rate (Statistics South Africa, 2008b). Table 5.1 contains a brief summary of the different features of the IES 2005/06.

Table 5.1: Brief summary of the Income and Expenditure Survey, 2005/2006

Feature		IES 2005/2006
Primary sampling units (PSUs)		3,000 PSUs
Sample size		24 000 dwelling units
Methodology		Diary and recall methods
Main questionnaire		One questionnaire (five interviews – i.e. five modules)
Diaries		Four weekly diaries
Expenditure data collection approach	Goods	Acquisition approach (i.e. the total value of all goods and services acquired, whether consumed or not, during a given period, whether or not they were wholly or partly paid for during the period of collection)
	Services	Payment approach
	Own production	Consumption approach
Survey period		One year (September 2005 to August 2006)
Reference period: Food expenditure		September 2005 to August 2006
Visits per household		Minimum of six visits
Classification of expenditure items		Classification of Individual Consumption According to Purpose (COICOP)

Source: Statistics South Africa (2008d)

The IES 2005/06 is different from the previous rounds that used only the recall method (i.e. one month for frequent spending and 12 months for spending on durables). The IES 2005/06 uses a combination of the recall method (through the main questionnaire) and the diary method for the first time in South Africa (Statistics South Africa, 2008d). The main questionnaire was divided into five modules and a new module is administered to each selected household in each successive week for five weeks. Households report their acquisition of all durable and semi-durable items over the 11 months preceding the survey month (Statistics South Africa, 2008d). In addition to the main questionnaire, one diary is given to each household in each successive week for four weeks to record daily acquisition and spending⁴⁵. The completed diaries were collected during regular visits when the main questionnaire is administered. Reported expenditures were annualised using the figures from both the main questionnaire and the diaries. For non-durable items (e.g. expenditure on food and non-alcoholic beverages, personal care and alcoholic beverages and tobacco) only the diary was used so that annualised expenditure was obtained by inflating monthly expenditure by a fixed multiplicative scalar of 12. Expenditures on durable and semi-durable items were collected using both the diary and the main questionnaire. Annualised figures were obtained by adding

⁴⁵ Households that are unable to complete the questionnaire were assisted by fieldworkers through regular visits.

the reported expenditure from the one month diary to those from the main questionnaire based on an 11 month recall period. Spending on services was collected using only the main questionnaire and with a recall period of 12 months. Therefore these represent annual figures and no further adjustments were needed.

For this study, the relevant information that will be extracted from the 2005/06 IES dataset includes household expenditure on goods and services, household consumption, household income and household composition and other related variables. These will be used to further extract information on household tax and health care payments. Specifically each household's contribution to general taxes (personal income tax, corporate income tax, value-added tax, fuel levy and excise taxes) allocated to the health sector of the economy and other contributions in the form of direct out-of-pocket payments and private insurance contributions (both employers' and employees' contributions) is extracted. Detailed explanation of the procedure for extraction is contained in the sections that follow⁴⁶.

5.3 MEASURING HOUSEHOLD INCOME

5.3.1 Alternative measures

The choice of an appropriate measure of ability to pay (ATP) for distributional analysis has been a topic of debate among economists (Sen, 1984, Musgrave, 1985, Blundell and Preston, 1995, Sahn and Stifel, 2003). In developed countries, apart from a few studies that use expenditure (Lancaster *et al.*, 1999), income is the preferred measure. In developing countries however expenditure or consumption is usually the preferred measure (Younger *et al.*, 1999, Sahn and Stifel, 2003, O'Donnell *et al.*, 2008a). The choice of expenditure over income in developing countries is motivated by the difficulties inherent in measuring income, "the seasonal variability in such earnings, and the large shares of income ... from self-employment both in and outside of agriculture" (Sahn and Stifel, 2003 p.464) which may not be recorded by households as income (O'Donnell *et al.*, 2008b). Income may also be concealed to evade taxation (Lancaster *et al.*, 1999). Other measures such as asset indices, and some other forms of composite measure of socio-economic status or community self-knowledge and self-rating (Filmer and Pritchett, 2001, Sahn and Stifel, 2003, Ichoku, 2011) have been proposed for developing countries but these are of limited use in distributional analysis of taxes or health

⁴⁶ See Table 5.5 for a brief summary of the procedure.

care payments (O'Donnell *et al.*, 2008b). Economic theory suggests that consumption is a better measure of long term welfare level than current income (Blundell and Preston, 1995). This is because of the possibility of smoothing of consumption while income may fluctuate with economic fortunes, activities and events.

In the context of tax incidence analysis, Bird and De Wulf (1973) discusses various measures of income that are often used. He notes that some studies use tax return data, some use the personal income or disposable income, others use income reported in household surveys without adjustments, while others make different assumptions on income distribution and obtain a 'rough' distribution of income based on different income classes. In this thesis, household consumption expenditure defined as the "[f]inal use of goods and services, excluding the intermediate use of some goods and services in the production of others" (O'Donnell *et al.*, 2008b p.70) is used as a proxy of household income. Specifically this was constructed as the summation of household spending on food items and non-food items including imputed values for home production and housing. In addition it was ensured that this includes all health care payments and is gross of direct taxes. This is because "[i]f one wishes to make an inference about the distributional impact of health finance..., then the measure of ATP should be gross of all health care, tax, and social insurance payments" (O'Donnell *et al.*, 2008b p. 188).

5.3.2 Adjusting income for underreporting

Measuring income in this fashion does not address the issue of underreporting that is common in household surveys (Robilliard and Robinson, 2003). Even when the household sampling weights are applied to aggregate the estimates of expenditure and income to obtain national aggregates they often do not tally (Deaton, 1997, Robilliard and Robinson, 2003). This has also been recognised with the South African IES 2005/06 (Statistics South Africa, 2008a, c). On this issue Statistics South Africa notes that "there was *clear* and *systematic* under-reporting in the IES" (Statistics South Africa, 2008a p.5 [emphasis added]) when expenditures in the IES are compared with aggregate figures. The underreporting was clear in the food and non-alcoholic beverages category. When compared with the IES conducted in 1995 and 2000, the share of food and non-alcoholic beverages expenditure in total expenditure fell from 28.5% (27.4%) in 1995 (2000) to 16.6% in 2005/06. Statistics South Africa attributed this to differences in survey methodology (Statistics South Africa, 2008c) particularly the use of diary method in the 2005/06 round. There were also some signs of underreporting in the alcoholic

beverages and tobacco category and other high-frequency items. Though there was underreporting, the relative shares of expenditure on the components that make up food and non-alcoholic beverages expenditure in the IES 2005/06 were similar to those recorded in the previous rounds. For example “bread and cereals ranged between 22.4% (1995) and 27% (2005/2006); meat ranged between 26.8% (2000) and 28.4% (1995); and vegetables ranged between 9.3% (1995) and 10.4% (2005/2006)” (Statistics South Africa, 2008c p.29). However it was remarked that “[i]nternational comparisons indicate that the proportion of consumption expenditure devoted to food in IES 2005/2006 is low for a country at South Africa’s level of development” (Statistics South Africa, 2008c p.17). This necessitates upward adjustment of this expenditure category. In response to this, and to confirm the underreporting, Statistics South Africa conducted a micro follow-up survey on 651 households using both the diary and the recall methods and it was found that “on average, expenditure by recall was 45% higher than expenditure by diary in the case of food and non-alcoholic beverages, and 212% higher in the case of personal care.” This supplementary evidence therefore strongly supports “an upward revision in the rand value of food and non-alcoholic beverages ... reported in IES 2005/2006” (Statistics South Africa, 2008c pp. 27-28). To adjust the IES 2005/2006, Statistics South Africa suggests the use of new CPI weights. They note that researchers may wish to take the CPI weights into account in analysis of the IES 2005/2006 data (Statistics South Africa, 2008c).

Very few studies have used the IES 2005/2006 in South Africa. One of such is Oosthuizen (2008). In adjusting total expenditure, because of the difficulties in determining the exact extent to which food expenditure items were underreported, Oosthuizen assumed a 20% under-reporting across all food and non-alcoholic beverages in his analysis of poverty lines in South Africa. Oosthuizen (2008) further noted that the 20% adjustment still underrepresented food and non-alcoholic beverages expenditure in total expenditure. This underrepresentation was also confirmed in this thesis. This approach, though ad-hoc, has been noted in Robilliard and Robinson (2003)⁴⁷. The need for adjusted expenditure is imperative in this thesis because the extracted taxes and health care financing from the IES were adjusted to tally with their *global figures*⁴⁸. The use of a uniform inflating proportion for food and non-alcoholic

⁴⁷ Robilliard and Robinson (2003) discusses the use of another methods – cross entropy in the case of reconciling household surveys with national accounts figures.

⁴⁸ See the next section for details on how these adjustments were performed and the assumptions made.

beverages may be supported by the fact that the relative share of different food and non-alcoholic beverages expenditure items in total food and non-alcoholic expenditure is similar across the 1995, 2000 and 2005/06 IES⁴⁹ (Statistics South Africa, 2008c). The proportion of food expenditure in total consumption expenditure in previous IES rounds (i.e. about 28%) was used to adjust total food expenditure in a proportional fashion across households. This adjustment ensured that expenditure on food and non-alcoholic beverages account for 28% of total household consumption expenditure.

Based on this, the new aggregate household consumption expenditure (y_2) was generated as:

$$y_2 = (y_1 - x) / 0.72 \quad (5.1)$$

where y_1 is the initial unadjusted aggregate household consumption expenditure and x is the initial unadjusted total food and non-alcoholic beverages expenditure. Each household's initial food/non-alcoholic beverages expenditure (x) was adjusted accordingly to satisfy the aggregate constraint $x_2 = 0.28y_2$ where x_2 is the new adjusted total food/non-alcoholic beverages expenditure.

5.3.3 Adjusting income using adult equivalence scales

Household income used as a measure to assess inequality and progressivity needs to further account for household size and composition (Coulter *et al.*, 1992, Deaton, 1997) in moving from total household to individual welfare. This requires some form of household equivalence scale measures. There are controversies on household equivalence scales and the most appropriate way to measure them (Coulter *et al.*, 1992, Banks and Johnson, 1994, Lancaster *et al.*, 1999). However, “[h]ousehold equivalence scales measure the relative income needs of households of different sizes and composition.... they answer the question of how much income different households would need to attain the same welfare level” (Nelson, 1993 p.471). This relates to the case that an additional household member may raise the cost of maintaining household welfare but proportionately less than the increase in household size (Glewwe, 1991). The easiest form of accounting for household size is the per capita approach.

⁴⁹ The assumption here is that these ratios remain fairly similar across years such that they are all underreported by a fairly similar magnitude. Using different inflating scalar would alter these ratios.

This is considered as naïve because it assigns to every household member the same welfare value irrespective of other differences. This has the potential to understate the welfare level of larger households and overstate that of smaller households because larger households tend to have more children (Nelson, 1993). This simplistic model of an equivalence scale follows the summary of the review by Buhmann *et al.* (1988) that an equivalence scale can be summarised as:

$$AE(N, \theta) = N^\theta \quad (5.2)$$

where N is the household size (addition of adults and children) and $\theta \in [0,1]$ is a parameter that measures need elasticity. When $\theta = 1$ we have the case of per capita estimates while $\theta = 0$ implies that differences in household size is irrelevant in the analysis as $AE(N, 0) = 1 \quad \forall N$. Higher values of θ imply decreases in economies of scale. Overall this is a simplistic and naïve assessment of equivalence scales that does not account even for differences in the composition of households and disregards intra-household decision-making and resource allocation.

Different countries have used different scales. The OECD for instance assigns the value of 1 to the first adult member, 0.7 to the second adult member and 0.5 to children (Gustafsson and Uusitalo, 1990, Ringen, 1991, Wagstaff *et al.*, 1999a).

In the literature, there are arguments as to whether inequality and poverty estimates could be sensitive to the choice of adult equivalence scales (Coulter *et al.*, 1992, Nelson, 1993, Jenkins and Cowell, 1994) or whether these differences could be partly associated with the data structure (Banks and Johnson, 1994). While it is likely that both of these effects may impact on inequality and progressivity estimates, in recent times there has been increased awareness and recognition in the literature that household consumption varies across household members in such a way that the consumption of a child is considered to be worth at most that of an adult member. Others even discuss this with respect to the sex of the household member (Buhmann *et al.*, 1988, Lancaster *et al.*, 1999) or differences in the age of children (McClements, 1977) or an income-dependent scale that accounts for whether the household is poor or rich (Aaberge and Melby, 1998). These are intra-household resource allocation issues. The main issue that is however largely left unresolved relates to how much less to weigh the consumption of non-adults relative to that of adults in generating household welfare levels. Understanding this issue

requires that we answer some questions such as: How is intra-household resource allocation conducted across households? Is such an allocation formula or pattern invariant across households? Because the answers to some of these questions are normative, the choice of an appropriate equivalence scale is still challenging (Deaton, 1997, Duclos and Araar, 2006) and as such many studies still adopt the naïve and simplistic per capita approach that follows from Equation (5.2).

Given that there is already a recognition of the effect of the choice of equivalence scales on the outcome measure of interest, sensitivity analyses are used to assess the extent to which the results remain robust to changes in the assumptions within reasonable limits (Lancaster *et al*, 1999).

As a result of these considerations, household income in this study is equivalised by the adult equivalence scale (Deaton, 1997) represented as:

$$AE = (n_A + \alpha n_K)^\theta \text{ for } 0 < \alpha \leq 1; 0 < \theta \leq 1 \quad (5.3)$$

where n_A is the number of adults in the household; n_K is the number of children (less than 15 years of age), α is the cost of children (a measure of the weight accorded to children relative to that accorded to adults) (Banks and Johnson, 1994, Citro and Michael, 1995) and θ represents a measure of economies of scale.

As noted above the choice of ‘appropriate’ parameters is fraught with practical and empirical challenges (Deaton, 1997). When the value of $\alpha = \theta = 1$, equation (5.3) can be interpreted as the total household size and deflating any welfare measure by this would yield per capita estimates (compare with $\theta = 1$ in (5.2)). Similarly as $\alpha \rightarrow 1$ the consumption of a child is assumed to be almost equivalent to that of an adult while $\theta \rightarrow 1$ indicates the absence of economies of scale such that larger households, on average, do not live more cheaply than smaller households. Deaton and Zaidi (2002) suggest that, for developing countries, the value of θ should lie between 0.75 and 1 while the value of α should lie between 0.3 and 0.5. This is because of the relative importance of food in total consumption, and the limited scope for economies of scale. It is important to note also that Woolard and Leibbrandt (1999) in their assessment of poverty in South Africa, have shown that their results are almost robust to the choice of parameters within these limits. In this thesis, unless otherwise stated, household per adult equivalent income will be computed using the values $\alpha = 0.5$ and $\theta = 0.75$ based on

(Deaton and Zaidi, 2002). The robustness of the results of the progressivity, and redistributive impact of health care financing to the choice of different scales within the limits noted in Deaton and Zaidi (2002) was later assessed.

5.4 INCIDENCE ASSUMPTIONS AND PROCEDURE FOR EXTRACTING HOUSEHOLD HEALTH CARE FINANCING CONTRIBUTIONS

As previously noted in South Africa health care is financed from different sources. The major sources include direct out-of-pocket payments, private health insurance contributions and general taxes. In 2005/2006 these make up about 14%, 45% and 40% of total health care financing respectively. In order to obtain household contributions to these sources the information mainly on reported income and expenditure contained in the IES 2005/2006 were used. In order to do this there was a need to make tenable assumptions on whom ultimately bears the incidence of the payments. The assumptions, including how these components were extracted, are provided in detail below⁵⁰.

Contributions to private *voluntary* health insurance (called medical schemes in South Africa) are, for most employed individuals, shared between employees and employers. Internationally it is recognised that the extent to which an employer would pass the burden of contributions unto employees depends on the elasticities of demand and supply of labour and other labour market factors (O'Donnell *et al.*, 2008b). While the supply of unskilled and informal labour is relatively wage elastic in South Africa, the supply of skilled and semi-skilled labour is relatively wage inelastic (de Wet and van Heerden, 2003). Those that purchase medical scheme cover are largely the formally employed skilled and semi-skilled workers who can afford to pay premiums. Frequently, enrolment is a condition of employment rather than aversion to risk (Ataguba and Goudge, 2012). Though there are recent efforts to provide health insurance cover for government employees, especially those in lower cadres through the Government Employees Medical Scheme (GEMS)⁵¹ (McLeod and Ramjee, 2007), enrolment was not widespread during the time the IES 2005/06 data was collected and as such members of medical schemes are largely skilled and semi-skilled workers with formal employment. In this

⁵⁰ A summary of this is also contained in Table 5.5.

⁵¹ GEMS was only registered in 2005 and became operational in 2006 (McLeod and Ramjee, 2007).

study therefore the assumption is that because the supply of labour (skilled and semi-skilled) is relatively inelastic, the possibility of shifting the burden unto the employee is very high. It was assumed that the entire burden is borne by the employee. In practice the basis for this assumption is that although employers may claim that they provide employees additional benefits related to the contributions they make on their behalf, the reality is that most employers in South Africa operate on a total 'cost to company' basis. Therefore they are able to shift their part of the contribution unto employees by adjusting the salary component that accrues to the employee. This is so because when employees choose a low cost medical scheme option, their salary component increases accordingly (Borghi *et al.*, 2009). This assumption is also in line with other international studies (see for example Wagstaff and van Doorslaer, 1992, 1997, Wagstaff *et al.*, 1999a, Bilger, 2008, O'Donnell *et al.*, 2008a). Therefore, for each household, the sum of employers' and employees' contributions is taken to represent the burden that the household bears.

The assumption about the distribution of out-of-pocket payments burden is relatively straightforward. These payments are assumed to be borne entirely by the consumer. This assumption has also been used internationally in the assessment of equity in health care financing (Castano *et al.*, 2002, Bilger, 2008, Yu *et al.*, 2008). Out-of-pocket payments were obtained as the sum of money spent on co-payments, medical services (including consultation with general practitioners, specialists, etc.), dental services, medical analysis laboratories and x-ray services, services of medical auxiliaries, hospital service fees, therapeutic appliances and equipment, medicines and pharmacy fees, traditional healers' fees, and other related medical products and service fees that are not reimbursed by any insurance scheme.

In South Africa while general taxes are not collected for the sole purpose of financing health care, a proportion of the revenue is allocated to the health sector (11.55% in 2006) (National Treasury, 2007). This is the proportion that is of relevance in this thesis. Because tax revenue forms an integrated pool of funds, a tenable assumption is that an equal proportion (i.e. 11.55%) of each tax category is allocated to the health sector. The following categories of taxes that when combined account for over 90% of total tax receipts in South Africa: personal income tax, corporate income tax, value-added tax, excise taxes, and fuel levy were considered.

The cardinal rule in the assessment of incidence is that "only individuals can bear the burden of taxation and that all tax burdens should be traced back to individuals." (Auerbach, 2006

p.2). Therefore assumptions are required in order to trace tax liability back to individuals. Individuals here could represent consumers, income recipients, labour and capital owners, etc. For the general taxes in South Africa, several assumptions on tax shifting or incidence are made depending on the type of tax involved. The process can be very complex and uncertain (Musgrave, 1985) especially for indirect taxes and corporate income tax (Musgrave, 1985, Kotlikoff and Summers, 1987, Martinez-Vazquez, 2001, Blackman *et al.*, 2010) as this depends on the conditions of demand and supply in different markets (Musgrave, 1985). The regressivity or progressivity of a tax system could be altered greatly by changing the assumptions made about tax shifting (Shah and Whalley, 1991). A classic assumption that most tax incidence studies, especially in developed economies, have adopted is based on Pechman and Okner (1974). Here income taxes are assumed to be borne by the income recipient while indirect taxes are shifted forward unto consumers. Tax incidence analyses have also been conducted under a general equilibrium framework (Harberger, 1962, McLure, 1975, Fullerton and Heutel, 2007). Under this framework changes in consumer and producer prices of all goods and services in the economy that are associated with the imposition of tax are used to assess the changes in consumers' welfare that results from such changes (Blackman *et al.*, 2010). In this thesis a partial equilibrium approach to incidence analysis is used with several assumptions on tax shifting. These assumptions are extensively discussed below.

In general, it is natural to assume that employees or income recipients bear the burden of personal income tax (Shah and Whalley, 1991). This assumption generally leads to progressive incidence largely because personal income tax rates (average or marginal) are structured progressively. In this thesis, the distribution of the burden of personal income tax is assumed to rest with the employee or the recipient of income. To extract an individual's total personal income tax liability two variables were considered from the IES - reported pay-as-you-earn (PAYE) and standard income tax on employees (SITE) - the sum of which should normally equal an individual's total personal income tax liability. From the IES, these variables were only recorded at household level and not at individual level. Again there was an indication of gross underreporting when compared with the figure on total revenue from this tax reported by the National Treasury. For this reason, personal income tax liability was computed through gross and taxable income calculated for all adult individuals from their reported income. Following this approach the total personal income tax liability extracted came up to about 90% of that reported by Treasury. Based on the components of income tax in South Africa, gross income was computed as the sum of salaries and wages received within the period of

assessment (this includes overtime, bonuses, cash allowances in respect of transport, clothing, and so on); income from business or professional practice/activities of farming; income from letting of fixed property; royalties; interests received and/or accrued on deposits, loans and savings certificates; dividends received on shares; and receipt of pension. To generate taxable income it was taken into account that natural persons under 65 years (65 years or over) of age are exempt from tax on interest and dividends up to R15,000 (R 22,000) per annum. Foreign dividends have a very low exemption limit but it was difficult to distinguish these from the dataset⁵². From the taxable income, the 2005/2006 tax table contained in Table 5.2 was applied using the tax thresholds and accounting for tax rebates. A household's total personal income tax liability was computed by summing up all tax liabilities of adult individuals in the household that are subject to personal income tax.

Table 5.2: Income tax rates, rebates, and thresholds in South Africa (2005/06 assessment year)

Taxable income (South African Rand)†	Rates of Tax (South African Rand)
< 80 001	18% for each R1
80 001 - 130 000	14 400 + 25% of the amount above 80 000
130 001 - 180 000	26 900 + 30% of the amount above 130 000
180 001 - 230 000	41 900 + 35% of the amount above 180 000
230 001 - 300 000	59 400 + 38% of the amount above 230 000
> 300 000	86 000 + 40% of the amount above 300 000
Rebates (Individuals only)	
Under 65 years (primary)	R6 300
65 years and older (secondary)	R4 500
Tax Thresholds	
Under 65 years	R35 000
65 years and older	R60 000

† note that the nominal exchange rate in 2006 was \$1 = R6.5

Source: National Treasury (2006)

While there is fairly little disagreement with the incidence assumptions of personal income tax, major areas of disagreements relate to some indirect taxes and corporate income tax. Corporate income tax is generally regarded as one of the most difficult taxes to allocate (Auerbach, 2006). It may be shifted 'forward' to consumers in the form of higher prices with a resulting regressive pattern or may sometimes be shifted 'backward' to recipients of capital income, in which case the tax is progressive (Shah and Whalley, 1991 p.536). Such shifts generally alter prices and returns on factors of production (Kotlikoff and Summers, 1987,

⁵² Also contributions to medical schemes should normally be deducted from gross income before applying the tax rates. However the dataset does not allow for such deductions. This is because contributions to medical aid/schemes were collected at the household level and not at the individual level. This may likely inflate the personal income tax liability computed for each individual.

Martinez-Vazquez, 2001). While it is relatively easier to assume that the statutory burden is borne by the firm upon which it is initially levied, it is very difficult to conjecture the distribution of the final economic incidence.

There are controversies on the extent to which the tax can be shifted forward unto consumers or backwards on capital owners or on labour (Musgrave, 1985). These controversies sometimes mean that some tax incidence studies do not include it or are silent on its treatment (van Doorslaer *et al.*, 1999, Younger *et al.*, 1999, Simkins *et al.*, 2000, Auerbach, 2006). To obviate this, Harberger (1962) was the first to use general equilibrium analysis to study the incidence of corporate income tax in the context of the United States. While it was generally believed that the bulk of corporate income tax burden is shifted forward unto consumers, Harberger (1962) though assuming that corporate income tax is the only tax in the system, showed that in the context of the United States capital seems to bear the bulk of the burden. The general equilibrium approach has continued to gain considerable attention among economists. However it relies on multiplicities of assumptions about the functioning of the economy that could also alter results. McLure (1975 p.128), while advocating for the use of general equilibrium models, argues that the standard analysis as opposed to general equilibrium analysis may be adequate and prove to be “extremely useful and illuminating under the proper circumstances, ... if handled with extreme care.” The same can be said of the general equilibrium analysis.

The general assumptions that follow from the traditional approach are related to the pricing policies that exist across sectors of the economy. In competitive markets there is a greater possibility that corporate income tax would be shifted more unto capital income (e.g. shareholders, contributors to pension funds, etc.) while consumers may bear the greater burden in the case of monopoly or oligopoly (see for instance Musgrave *et al.*, 1951). This requires an understanding of the underlying market structure in different sectors of the economy. Most studies in developing countries usually assume that 50% of the burden of corporate income tax is shifted forward unto consumers while 50% is shifted backwards unto capital; or different complementary shares (Huang, 1976). These assumptions are used mainly because of the difficulty in specifying the exact extent of pricing in different markets. A very useful measure of business concentration used in the literature is the Hirschman-Herfindahl Index (HHI) (Rosenbluth, 1955, Hirschman, 1964). This index indicates the market structure of industries in a country.

Specifically **HHI** indicates how far the industry is from perfect competition or monopoly. Theoretically the index ranges from 0 to 1. Zero indicates perfect competition while 1 indicates monopoly and it can be generally characterised as follows: (0 – 0.25) for weak, (0.25 – 0.50) for moderate, (0.5 – 0.75) for fairly strong and (0.75 – 1.00) for very strong with regard to how close the industry is to being a monopoly (Djolv, 2009). To date no study on corporate income tax has used this to underlie assumptions about the distribution of tax burden.

In the case of South Africa the **Herfindahl** index has recently been computed as contained in Table 5.3. This shows that between 2003 and 2006 there are general tendencies that manufacturing industries are characteristically leaning towards competition or oligopoly. For instance in 2006 the index averaged 0.37 with the market for beverages recording a value greater than 0.6. This trend is fairly consistent across the years. The implication of this is that part of the burden of corporate income tax is likely shifted forward unto consumers and others to capital owners. While this can be deduced, the exact extent of the split is still difficult to ascertain. Further, company income taxation in South Africa is very complex. Smaller companies are taxed differently from larger ones. There is also a special type of company income tax called the *Secondary Tax on Companies (STC)*. This is based on net dividends (dividends declared less all dividends receivable during a dividends cycle). In general, the normal tax rate on companies is 29% of declared incomes and the *STC* is 10% of the net dividends. Micro or small businesses have graduated tax rates ranging from about 1% to 7%. Because the *STC* component works via dividends, and the general company tax works via declared incomes, again there is an indication that the burden could be shared by shareholders and consumers. Given these complications and the average **Herfindahl** index value of 0.37 (in 2006), it was assumed that the split between capital owners (shareholders, contributors to pension, etc.) and consumers would be in the neighbourhood of 40% and 60%⁵³. Capital owners were identified as those who report dividends and contributions to investment. Reported corporate income tax revenue obtained from National Treasury was then apportioned to consumers and capital owners accordingly.

⁵³ Unless otherwise specified, the 40:60 split respectively between consumers and capital owners was used in this thesis. Though there is no direct mathematical relationship between the **Herfindahl** index and the proportions used to split the burden, the 40:60 split was used because the **Herfindahl** index value of 0.37 leans more towards competition than monopoly. In addition to this, sensitivity analyses were conducted to assess the impact of the choice of the split on robustness of the results. The two extreme conditions (i.e. where the entire burden is borne by consumers and where the entire burden is borne by capital owners) were included in the sensitivity analysis.

Table 5.3: Herfindahl Index, Ratio for South Africa (2003 - 2006)

Category	2003	2004	2005	2006
Food products	0.4449	0.4503	0.4479	0.4476
Beverages	0.5843	0.6212	0.6076	0.6145
Tobacco	0.6476	0.2442	0.3324	0.3693
Textiles	0.2413	0.1761	0.1432	0.1835
Wearing apparel, except footwear	0.2909	0.2929	0.2496	0.2924
Leather and fur products	0.2066	0.3640	0.3149	0.3035
Footwear, except rubber or plastic	0.1920	0.2004	0.2046	0.3341
Wood products, except furniture	0.3861	0.1946	0.1626	0.3300
Furniture and fixtures, excluding metal	0.2633	0.2076	0.2062	0.2264
Paper and products	0.4950	0.4407	0.5356	0.5158
Printing and publishing	0.3714	0.3004	0.2982	0.3283
Industrial chemicals	0.5372	0.5153	0.4988	0.4969
Other chemicals	0.4014	0.3877	0.3407	0.3664
Petroleum and related products	0.2914	0.3665	0.2569	0.2106
rubber products	0.3835	0.5053	0.4268	0.4594
Plastic products	0.3574	0.2560	0.2383	0.2031
Non-metallic mineral products	0.4364	0.3514	0.3379	0.3251
Glass and products	0.5689	0.5442	0.5309	0.5483
Iron and steel	0.5797	0.6425	0.6051	0.6006
Non-ferrous metals	0.5495	0.4997	0.4881	0.5261
Fabricated metal products	0.2966	0.3270	0.2695	0.1852
Non-electrical products	0.3622	0.3897	0.3290	0.2388
Electrical products	0.3557	0.2457	0.2350	0.2561
Transport equipment	0.5684	0.5389	0.5114	0.5526
Professional and scientific equipment	0.3832	0.1360	0.1503	0.2141
Other manufacturing industries	0.4408	0.3951	0.2266	0.3252
Manufacturing (average)	0.4122	0.3716	0.3462	0.3681

Source: Djolov (2009).

Value-added tax was introduced in South Africa in 1991 to replace the general sales tax and the rate increased from an initial 10% to 14% (in 1993) on all standard rated goods and services including certain imports (Go *et al.*, 2005). VAT is often regarded as a regressive form of taxation and this has been shown in the context of South Africa (Fourie and Owen, 1993, Sahn and Stifel, 2003, Go *et al.*, 2005). Because VAT is a tax, not on intermediate goods but on the final value-added, the most common assumption is that the burden is passed onto the final consumer. This assumption is used in this thesis to allocate the burden of VAT. Certain category of goods and services are VAT-exempt or zero rated (i.e. final consumers do not explicitly pay VAT on them). Domestic VAT-exempt goods and services include residential accommodation rentals, educational services, medical services at public facilities, fee-free financial services, and public road and rail services while zero-rated goods and services include major items consumed by the poor such as brown bread, lentils, beans and other legumes,

maize meal, rice, fresh vegetables, eggs, milk and fresh fruits. In South Africa total VAT revenue is composed of domestic and customs related VAT. The domestic VAT relates to that described above. The customs-related VAT is imposed on certain goods imported into South Africa. This includes imports made by ordinary individuals, firms and organised groups. In the 2005/06 assessment year customs-related VAT accounted for about 30% of total VAT receipts. There are also items exempted from the customs-related VAT and these include mainly those imported by individuals such as wines, perfumes, cigarettes, and some used items subject to some quantity limits. Foreign nationals on short visits to South Africa can reclaim VAT payments on certain unused items.

A household's VAT liability was estimated from household reported expenditure on items that are standard rated excluding exempted and zero rated items. Basically to obtain the VAT liability of household i (VAT_i), household expenditure items that are standard rated are summed together ($\sum_j g_j$) then the VAT contribution is obtained as:

$$VAT_i = (0.14 / 1.14) \sum_j g_j = 0.123 \sum_j g_j \quad (5.4)$$

Excise taxes, excluding fuel levy, are imposed on certain locally manufactured goods and their imported equivalents. These items include tobacco products, liquor (spirits) products, traditional African beer, malt beer, wine and other fermented beverages⁵⁴. The major assumption is that excise taxes are shifted forward on to consumers in the form of higher prices (Shah and Whalley, 1991, Cook and Moore, 1994, Gemmell and Morrissey, 2005). There are disagreements on this (Shah and Whalley, 1991), and what happens in practice is likely to be in part a function of the existing market structure (Barde and Braathen, 2005). In this thesis it was assume that the burden of all the excise taxes is borne entirely by the consumers. This assumption has recently been used to assess the progressivity of alcohol taxes in South Africa (Ataguba, 2012a) and is consistent with that used internationally. Items subject to excise tax are taxed at different rates. In the 2005/06 assessment period the rates range from R0.08/litre of traditional beer, R1.68/litre for beer, R1.41-R3.88/litre of wine, to R50.42 per litre of absolute alcohol (i.e. R20.17/litre of spirits⁵⁵) (National Treasury, 2006). Cigarettes and related products are levied as a specific tax of 52% on the retail price of the product.

⁵⁴ Soft drinks were not included as excise tax on soft drinks in South Africa was removed in 2002.

⁵⁵ This is assuming an average alcohol content of 40% per volume.

Total Excise tax liability for household i (E_i) is calculated based on reported household expenditure on goods and services where these taxes are applicable:

$$E_i = \underbrace{\sum_j (c_j / p_j) \cdot r_j}_{\text{wine, beer, spirits}} + \underbrace{0.52 \sum_l (c_l)}_{\text{cigarette products}} \text{ for each } i \quad (5.5)$$

where c_j is the household expenditure on commodity type j ; p_j is the average price per litre of commodity type j ; r_j is the excise tax rate for commodity type j ; and c_l is the pre-VAT cost of the cigarette product. Average prices used were based on the Consumer Price Index (CPI). The actual prices used in the construction of the national and regional CPI for 2005/2006 are not available. These prices have been published regularly by Statistics South Africa until 2000. From 2001 only the adjusted price index is available. Prices for 2005/06 were obtained from the 2005/06 CPI and the actual prices reported in 2000.

Fuel taxes have both direct and indirect effects (Mabugu *et al.*, 2009, Blackman *et al.*, 2010) which is the product of linkages between sectors in the economy. The direct effect of fuel tax is borne by those who spend directly on fuel (e.g. diesel, gasoline, kerosene, etc.) while indirect effects come from those who either spend on commercial transportation or through the purchase of goods and services produced using some form of fuel energy as an intermediate input, or where intermediate inputs are produced using fuel energy. By spending on these goods and services, households are contributing to fuel tax revenue indirectly. Many studies on fuel tax incidence focus on the direct aspects and ignore the indirect effects (see for example Poterba, 1991, Walls and Hanson, 1999) primarily because of the difficulty in estimating the indirect effect through household consumption and public transportation (Blackman *et al.*, 2010). Others have attempted to include the indirect effects either through partial equilibrium analysis (Blackman *et al.*, 2010) or some form of general equilibrium framework (Mabugu *et al.*, 2009). In the case of South Africa, the fuel levy is paid by both direct (private transport owners) and indirect users (commercial transport users and consumers of manufactured products) of fuel (Mabugu *et al.*, 2009). While the burden of direct payments is concentrated on the richest population, indirect burdens are concentrated on the poorer groups (Mabugu *et al.*, 2009) that rely more on public transportation. In this thesis, for direct users it was assumed that the burden is passed onto the consumer. For commercial transport users, even though there are some forms of public transportation such as buses and trains that are subsidised by the government, it is assumed that commercial transport operators are able to pass this forward

unto the consumers in the form of higher transport fares. For the fuel levy component attributable to other indirect fuel users (i.e. through the consumption of manufactured goods), based on international practice, it is assumed that demand for most of these goods are relatively inelastic (Blackman *et al.*, 2010) hence the fuel tax is passed forward to consumers in the form of higher prices resulting from increased cost of production. Even in a general equilibrium framework (Mabugu *et al.*, 2009) a similar assumption is implicit in the Social Accounting Matrix used.

Fuel levy in South Africa is imposed on specific ‘fuel’ goods including petrol (lead or unleaded), illuminating kerosene, and diesel. Apart from international movements in oil price, fuel tax⁵⁶ is a major contributor to increases in the fuel pump price in South Africa (Mabugu *et al.*, 2009). In the 2005/06 assessment period the rates are: R1.27/litre for petrol and R1.11/litre for diesel.

Total fuel levy liability for a household is the sum of direct and indirect liabilities estimated from reported household expenditure. Total fuel levy liability by direct users is calculated directly from reported expenditure (see Blackman *et al.*, 2010 for a similar procedure) on fuel as:

$$F_i^a = \sum_j (c_j / p_j) \cdot r_j \text{ for each } i \quad (5.6)$$

where c_j is the household private expenditure on the fuel type j ; p_j is the price per litre of fuel type j ; and r_j is the fuel levy rate for fuel type j . Because the survey did not distinguish petrol from diesel, the average levy and price was used for r_j and p_j .

For commercial transport users, fuel levy was estimated accordingly as:

$$F_i^b = \sum h \cdot r (c / q^*) \text{ for each } i \quad (5.7)$$

where c is household i 's expenditure on commercial transportation; q^* is the average cost of transportation per kilometre (q) divided by four⁵⁷; h is the estimated rate of fuel consumption per km; and r is the average fuel levy rate of diesel and petrol. Two difficulties were

⁵⁶ It is important to note that the price of fuel is composed of many parts. Fuel levy is only one of such components.

⁵⁷ Here four passengers were assumed to be the average number of passengers travelling per kilometre.

encountered here. One is determining, from the dataset, the type of fuel used by every mode of transportation and the other is average fuel consumption rate of each mode of transportation. With regard to these and particularly the latter, the average fuel consumption rate available from an external study (Prozzi *et al.*, 2002) was used. The values of h and q are contained in Table 5.4.

Table 5.4: Average cost of transportation and fuel consumption

Parameter	Value
Public mode of transportation* (q)	180 cents/km (in 2006 prices)
Fuel consumption (average of all modes) [†] (h)	0.2124 litres/km

* Obtained from the prices used in computing consumer price indices.

[†] Obtained from Prozzi *et al.* (2002).

The fuel levy component attributable to other indirect fuel users (i.e. through the consumption of manufactured goods) was estimated as the difference between total fuel levy revenue reported by the National Treasury and the amount accounted for by direct and commercial transport users calculated from the IES dataset. This was allocated to households based on their relative share in expenditure on manufactured commodities.

A brief summary of the methods used in estimating each component of the health care financing system is contained in Table 5.5.

Table 5.5: Summary of the computation methodology for each component of health care funding

Component	Share in Total Health care financing	Rates*	Computation Technique
Taxes	40.41%		
Personal Income Tax	12.1%	<p>18-40% depending on income level plus flat rate contribution as a function of income.</p> <p>0% tax for incomes below R35,000 for individuals below 65 years and R60,000 for individuals 65 years or older.</p> <p>A rebate of R6,300 and R4,500 for individuals below 65 years and 65 years or older respectively.</p>	Apply the appropriate tax thresholds, tax rate and rebates to the gross income of individuals within each household within the taxable range
Corporate Income Tax	10.1%	Mainly 29%	Apportioning total corporate tax receipts reported by South African Treasury to households based on different assumptions of tax shifting in terms of percentage borne by shareholders (identified as those who report earning dividends) and the rest by households through consumption. Two extreme scenarios were initially considered – that shareholders/capital owners bear the entire burden of the tax, and that consumers bear the entire burden. Based on the reported Herfindahl index value of 0.37, the following were considered: (40%:60%), (50%:50%), and (60%:40%) as plausible split of the burden respectively between consumers and capital owners in each scenario. In the general results presented the 40%:60% split was used but also sensitivity analysis was performed to test the robustness of the results to changes in these assumptions. Within these neighbourhoods, the results were robust.
Value-Added Tax (VAT)	10.9%	14% on standard rated goods and services	The VAT rate is applied to expenditure on goods and services that are standard rated, i.e. excluding the zero-rated and exempted goods.
Fuel Levy	2%	R1.27/litre for petrol and R1.11/litre for diesel	Since fuel is consumed by households (personal or commercial transportation) as well as corporate or industrial users, estimation involved a process of generating the component attributable to commercial transport users, personal transport users and users in businesses. It was assumed that the fuel levy is shifted to consumers reporting expenditure on minibus taxis, buses and other types of commercial transport. Fuel tax accruing to businesses and corporate

Component	Share in Total Health care financing	Rates ^a	Computation Technique
Excise Tax	1.5%	52% of retail price for cigarettes R1.68/litre for beer R0.08/litre of traditional beer R1.41-R3.88/litre of wine R20.17/litre of spirits ^b	users is also assumed to be passed forward onto consumers. Because the component attributed to corporate or industrial users could not be directly estimated from the dataset, it was assumed that the difference between the fuel levy component accounted for by private and commercial transport users and that reported by National Treasury is attributable to industrial users. For cigarettes and related products, the tax rate was applied to the expenditure on cigarette products. For beer, wine and spirits, reported expenditure on these products was translated into estimated quantities (litres) using average retail prices; the rate per litre was then applied.
Others	3.7%	Includes taxes on property and unidentified levies, stamp duties and fines, air departure tax and skills development levy.	Not estimated
Health Insurance	45.27%		
Medical Schemes (Private)	45.27%	-	Expenditure on medical scheme premiums by households was combined with employers' contribution on behalf of members of the household.
Out-of-pocket (OOP) payment	14.32%		
OOP payments	14.32%	-	Household expenditure on medicines, consultations, treatments and procedures were summed.

^a This applies to the taxes in the 2005/2006 assessment year.

^b assuming 40% alcohol per volume

Note: IES – Income and Expenditure Survey.

Source(s): Many sources including National Treasury (2006) and McIntyre *et al.* (2007).

5.5 ADJUSTMENTS AND BENCHMARKING OF EXTRACTED FINANCING COMPONENTS

As noted earlier, only 11.55% of general taxes were allocated to the health sector and this is the only portion that was relevant in the analysis of tax-related health care financing. Before extracting this proportion, it is important that the dataset is benchmarked to reflect tax receipts during the assessment period. This method of adjustment has been outlined in Borghi *et al.* (2009). For the various components of tax, the aggregate tax extracted from the IES taking into account the sampling weights was compared with total tax receipts reported by National Treasury. In all cases the IES underreported total tax receipts. This was particularly the case for value-added tax. Extracted value-added tax amounted to about 40% of total VAT receipts indicated by National Treasury⁵⁸. Personal income tax and excise taxes were very close. Extracted personal income tax accounted for about 90% of the amount indicated by National Treasury while excise taxes accounted for over 94%. Private health insurance contributions were also very close to the aggregate figure. The aggregate figure was obtained from the annual report of the Council for Medical Schemes (Council for Medical Schemes, 2006). Over 83% of the national aggregate was extracted from the dataset. In all these cases to ensure that the dataset is benchmarked to the 2005/2006 assessment year, the various components were adjusted accordingly as:

$$T_{Adj}^k = \sum_i (D \cdot f_i + T_i) \text{ for each } k \quad (5.8)$$

where T_{Adj}^k is the total adjusted k payment variable (e.g. private insurance, personal income tax, etc.), D is the deficit to be allocated, f_i is household i 's weighted relative frequency of the payment variable T_i . The term $D \cdot f_i$ here is the adjustment or inflation factor.

While it is relatively easy to obtain aggregates for private insurance contributions from the Council for Medical Schemes and even taxes from National Treasury, it is difficult to obtain equivalent aggregates for out-of-pocket payments. This is because there is no entity that keeps account of such payments. To arrive at fairly accurate aggregate, information were drawn from different sources. If the sum of general taxes, private insurance and out-of-pocket payments

⁵⁸ Notice that the underreporting of food expenditure noted earlier imparted on the proportion of VAT liability extracted from the dataset. This is because VAT was computed on the unadjusted food expenditure.

account for close to 100% of total health care financing, based on their relative contributions, out-of-pocket payments can be computed as the residual. This 'residual' is taken as the global figure. To adjust total out-of-pocket payments, user fee revenue was obtained for all public health facilities across all provinces in South Africa and co-payments made by members of private insurance schemes. The total hospital user fees revenue was used as an estimated value of user fees received (i.e. out-of-pocket payments to public sector providers). All public health care out-of-pocket payments from the IES were then added up and apportioned to households the difference between that obtained from the IES and that obtained as user fees from public hospitals. The reported out-of-pocket payments by scheme members based on the IES was very close to that obtained from the administrators of the largest medical schemes in South Africa so there was no specific adjustment made for this component. Specifically it was observed that out-of-pocket payments to public service providers were underreported by 37.54% and this was used to adjust that component accordingly. To double-check, total out-of-pocket payments from the IES were obtained after this adjustment and compared with that based on the 'residual approach' and the figures were very similar. As a result no further adjustment was made.

Chapter 6

EMPIRICAL ASSESSMENT OF THE INCIDENCE OF HEALTH CARE FINANCING IN SOUTH AFRICA

6.1 INTRODUCTION

This chapter provides an in-depth empirical analysis of the incidence of health care financing in South Africa. Each source of health care financing and overall health care financing is investigated. The methodology used includes simple descriptive presentation of the incidence of health care payments, the Kakwani index of progressivity and stochastic dominance tests. Robustness and sensitivity analyses were also performed to assess the sensitivity of the results to assumptions made about the equivalence scale parameters, shifting assumptions about corporate income tax and choice of parameters of aversion to inequality. Results from this chapter are also discussed in the light of international evidence and health care financing challenges in South Africa. The discussions in this chapter centre on the ability to pay principle earlier introduced in Chapter 3. Relative progressivity is therefore assessed in relation to households' ATP.

6.2 OVERVIEW OF THE METHODOLOGY

Under the methodological review chapter (Chapter 4) the various methods that can be used to assess progressivity of health care financing are reviewed. Irrespective of the choice of method, three broad conclusions can be deduced from the assessment of progressivity. Health care financing could be progressive, regressive or proportional. In this chapter, to investigate the incidence of health care financing in South Africa, the first step was to divide the entire sample into five equivalised income quintiles and show each quintile's share of income spent on health care through taxes, out-of-pocket payments, private health insurance and overall health care

financing. This can be denoted as the quintile's payment ratios (an analogue of the average tax rate). If we denote each quintile's income and payment as x_i and T_i respectively, then progressivity can preliminarily be assessed by looking at how (T_i / x_i) varies with x_i . As illustrated in Equation (3.1) a positive (negative) slope would imply a progressive (regressive) relationship because payments are an increasing (decreasing) proportion of incomes. Secondly Kakwani indices were obtained for each of the financing mechanisms and overall health care financing. The Kakwani index was computed using DASP software⁵⁹ (Araar and Duclos, 2009a, b) as the difference between the concentration index of payment (C_T) and the Gini index of income distribution (G). As noted earlier when the concentration index of payment exceeds (is less than) the Gini index of income distribution, the payment schedule is described as progressive (regressive).

The Gini index was computed as:

$$G = 1 - (\hat{\xi} / \hat{\mu}) \quad (6.1)$$

where $\hat{\xi} = \sum_{i=1}^n ((V_i)^2 - (V_{i+1})^2) / (V_1)^2 x_i$; $V_i = \sum_{h=1}^n w_h$ and $x_1 \geq x_2 \geq \dots \geq x_{n-1} \geq x_n$. The vector $\mathbf{x} = [x_1, x_2, \dots, x_n]$ represents a vector of per adult equivalent incomes while the vector $\mathbf{w} = [w_1, w_2, \dots, w_n]$ represents the sampling weights. $\hat{\mu}$ is the weighted average of per adult equivalised incomes.

Analogously the concentration index for the health care payment variable T was computed as follows using the vector of per adult equivalent incomes as the ranking variable.

$$C_T = 1 - (\hat{\xi}_T / \hat{\mu}_T) \quad (6.2)$$

where $\hat{\mu}_T$ is the weighted average of T ; $\hat{\xi}_T = \sum_{i=1}^n ((V_i)^2 - (V_{i+1})^2) / (V_1)^2 t_i$; and $x_1 \geq x_2 \geq \dots \geq x_{n-1} \geq x_n$.

From equations (6.1) and (6.2) Kakwani index of progressivity (π_K) for the health care payment variable T is obtained as:

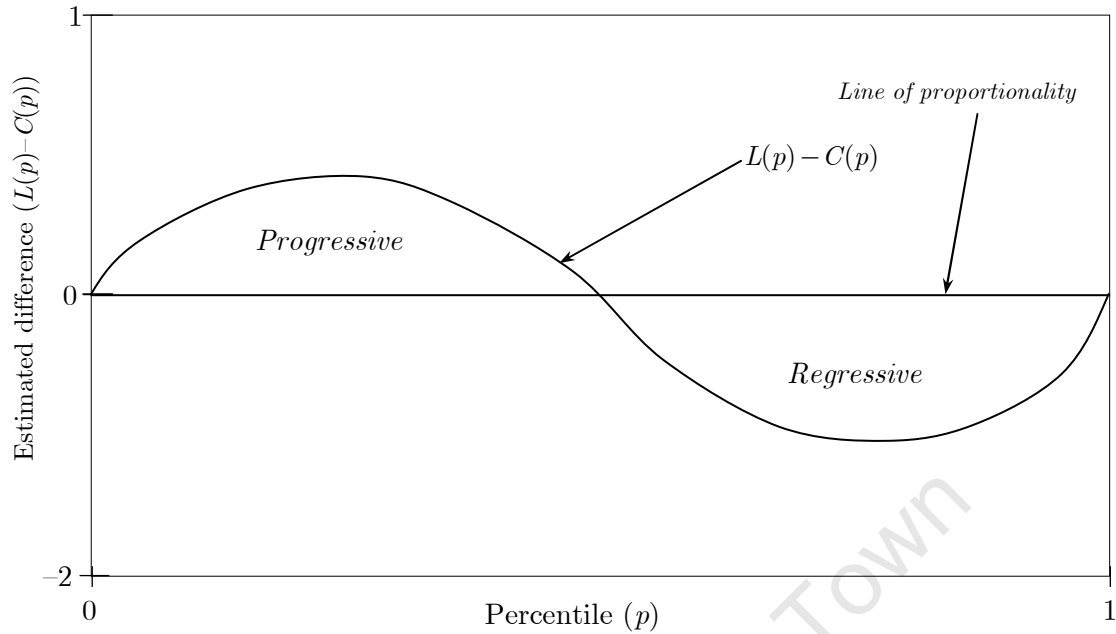
$$\pi_K = C_T - G = (\hat{\xi} / \hat{\mu}) - (\hat{\xi}_T / \hat{\mu}_T) \quad (6.3)$$

⁵⁹ DASP stands for Distributive Analysis Stata Package. This software is developed by Duclos and Araar and it is freely available (<http://dasp.ecn.ulaval.ca/index.html>).

where $\pi_K > 0$ represents a progressive health care financing mechanism; $\pi_K < 0$ a regressive mechanism; and $\pi_K = 0$ a proportional financing mechanism. The standard error for this index was obtained analytically using the DASP software (Araar and Duclos, 2009b) accounting for both the sampling design features of the IES 2005/06 survey (discussed in Chapter 5) and the dependence of the concentration and Gini indices obtained from the same dataset.

While the Kakwani index of progressivity is a summary index, it does not provide a full picture of the extent of progressivity across the entire income distribution percentiles. In order to show this the difference between the Lorenz curve and concentration curve coordinates was computed and plotted at different percentiles of per adult equivalent income distribution with their respective standard errors. This is the visual analogue of the dominance test between the concentration curve and the Lorenz curve (income distribution). As shown in Figure 6.1, when the difference between the Lorenz curve coordinate at percentile p (i.e. $L(p)$) and the concentration curve coordinate at the same percentile ($C(p)$) is positive (i.e. $L(p) > C(p)$) the financing mechanism is progressive at percentile p and the Lorenz curve dominates the concentration curve at the same percentile point. A regressive relationship occurs when $L(p) < C(p)$ such that the Lorenz curve is dominated by the concentration curve at percentile p . When both the concentration curve and the Lorenz curve coincide such that $L(p) = C(p)$, the relationship at percentile p is said to be proportional. Statistical significance of these differences across percentile points was assessed using DASP software (Araar and Duclos, 2009a). A 95% confidence interval was constructed and plotted around each progressivity curve.

Figure 6.1: Progressivity and regressivity illustrated using the progressivity curve



A formal test of stochastic dominance was also performed to determine whether the Lorenz curve dominates the concentration curve (i.e. a progressive relationship) or vice versa (i.e. a regressive relationship) (Bishop *et al.*, 1994, Davidson and Duclos, 1997, Sahn *et al.*, 2000, Duclos and Araar, 2006). We can state a sufficient condition for stochastic dominance of the concentration curve ($C_T(p)$) by the Lorenz curve ($L(p)$) as:

$$L(p_i) \geq C_T(p_i) \quad \forall p_i \in [0, F(\lambda^+)] \quad (6.4)$$

where $F(\lambda^+)$ is the cumulative distribution of individuals/households ranked by per adult equivalent income.

If the inequality in (6.4) holds up until $\lambda = +\infty$ then we can conclude that there is stochastic dominance between the Lorenz curve $L(p)$ and the concentration curve $C_T(p) \forall p_i \in [0, 1]$. Therefore the health care financing mechanism T is, weakly speaking, everywhere progressive. Strict progressivity will require that $L(p_i) > C_T(p_i) \quad \forall p_i \in [0, 1]$.

Generally, the same procedure can be used to test the stochastic dominance between two curves (e.g. two concentration curves). If we denote two curves A and B as C^A and C^B respectively then the distribution represented by C^A dominates that by C^B if $C^A(p_i) \geq C^B(p_i)$

$\forall p_i \in [0,1]$. This procedure imposes strict constraints on the assessment of stochastic dominance especially at the tails of the distribution (Sahn *et al.*, 2000) which could exhibit ‘noises’. In empirical application, therefore, equation (6.4) is applied on the percentile range $[0.05, 0.95]$.

I use the multiple comparison approach (MCA) to stochastic dominance to assess dominance between curves (Dardanoni and Forcina, 1999). The procedure has been applied in related research (see for example Wagstaff *et al.*, 1999b, O'Donnell *et al.*, 2008a, O'Donnell *et al.*, 2008b). Under this procedure p_i (percentile) in equation (6.4) is replaced with q_i (quantile points) and stochastic dominance is evaluated in the same manner. 19 equally spaced quantiles were used after restricting the percentile range to $[0.05, 0.95]$. However, in comparison to the strict criteria, using the MCA runs the risk of inferring dominance results too often.

As noted earlier, three broad health care financing mechanisms were considered – general taxes, out-of-pocket payments and private health insurance. For general taxes two broad sub-groups were created (i.e. direct and indirect taxes) that aggregates the components that they are made up of. Direct taxes comprise personal income tax and corporate income tax while indirect taxes comprise value-added tax, excise taxes and fuel levy. Chapter 5 (the detailed methodology chapter) contains an in-depth explanation on how these were estimated and extracted from the IES including all the assumptions used.

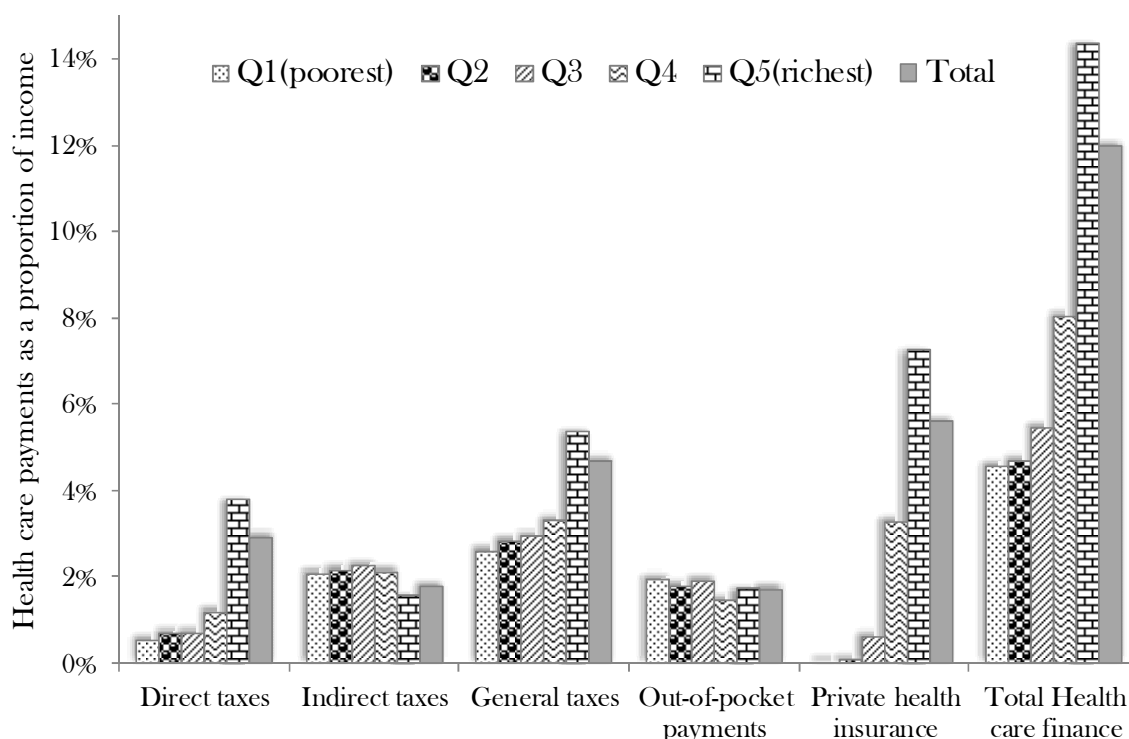
In the sections that follow the results of the incidence of health care financing in South Africa are presented. This also includes robustness and sensitivity of the results to the choice of various parameters.

6.3 EMPIRICAL RESULTS OF THE INCIDENCE OF HEALTH CARE FINANCING IN SOUTH AFRICA

The distribution of health care financing burden in South Africa as shown in Figure 6.2 indicates that direct taxes, private health insurance contributions and general taxes are progressive health care financing mechanisms while indirect taxes and out-of-pocket payments are regressive. It should be noted here that only the proportion of tax revenue that is allocated to the health sector (11.55%) is considered in the incidence analysis. Progressive financing mechanisms indicate that poorer quintiles spend less as a proportion of their income than

richer quintiles in financing health care. For example while the overall average percentage of income spent on health care via direct taxes is 3%, the poorest 20% of the population spends about 0.5% of their income compared to 4% spent by the richest 20%.

Figure 6.2: Health care payments as a proportion of income, South Africa 2005/2006



Note: Income and payments are expressed per adult equivalent

While the distribution of direct taxes is progressive, indirect taxes are regressive. Though on average 1.8% of total income is spent in indirect taxes that are allocated to health care, the bottom 20% of the population spends more than 2% of their income in indirect taxes for health care compared to 1.6% spent by the top 20%. The distributional pattern of the burden of indirect taxes across the other quintiles is less clear as it tends to be progressive up until the third quintile.

While health care financing via direct taxes generally looked progressive and indirect taxes looked regressive judging from the graphs in Figure 6.2, general taxes (a combination of direct and indirect taxes) is progressive. The bottom 20% of the population spends 2.6% of their income in general taxes allocated to health care compared to 5.4% by the top 20%. On average about 4.7% of incomes are spent in general taxes to finance health care in South Africa.

As depicted in Figure 6.2 out-of-pocket payments generally looked regressive. This is because the bottom 20% of the population spends more as a proportion of their income out-of-pocket

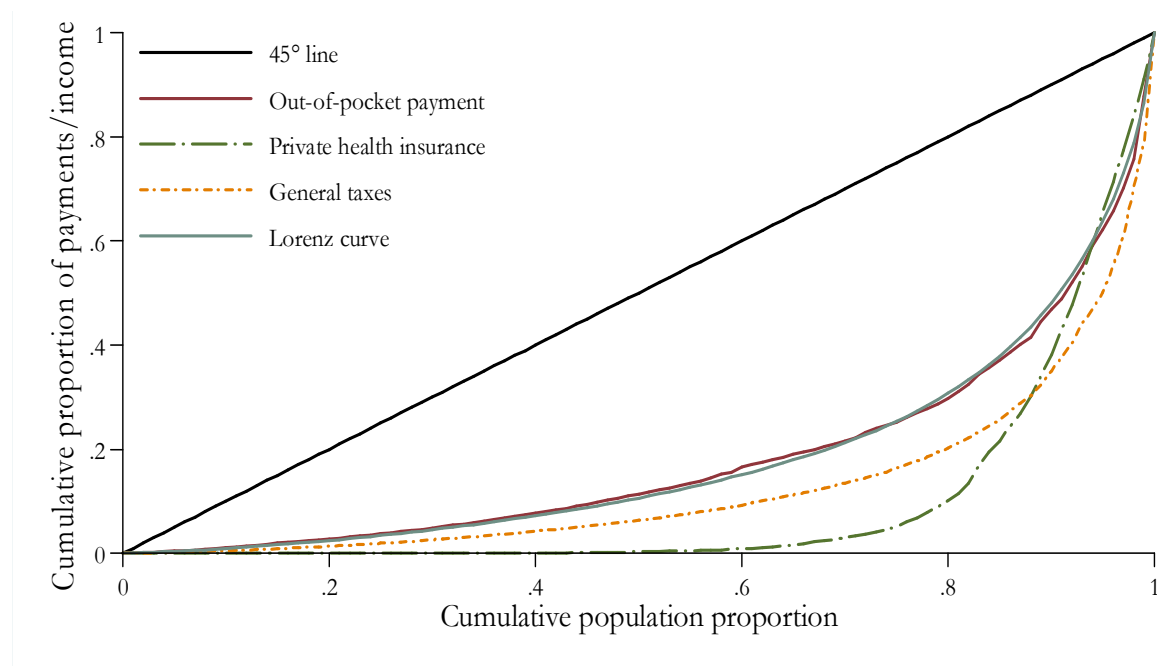
compared the top 20% of the population. The proportion spent by the top 20% and the overall average is very close.

On average about 5.6% of income in South Africa is spent financing health care through private health insurance contributions but this fraction is higher among the richest quintile (7.3%) compared to the poorest quintile (0.02%). This also reflects the distribution of households that purchase private health insurance cover in South Africa. From the IES 2005/2006 dataset, only about 16.5% of South Africans report membership of any private health insurance. Over 76% of this number are in the top 20% of the income distribution while <1% are in the bottom 20%.

Overall health care financing is progressive because private health insurance and general taxes are progressive and their combined share in total health care financing is greater than 80%. Figure 6.2 shows that on average 12% of income is spent financing health care in South Africa. The top 20% of the population spends about 14% of their income in financing health care compared to the bottom 20% that spends <5%. The second, third and fourth quintiles spend about 4.7%, 5.5% and 8% respectively in total health care financing. The bulk of the burden of total health care financing for the top 20% of the population is accounted for by private health insurance contributions. For this group private health insurance contributions account for over 50% of the burden of total health care payments compared to 37% and 12% by general taxes and out-of-pocket payments respectively. For the bottom 20% of the population general taxes (due to the regressivity of indirect taxes) accounts for over 57% of the total burden of health care financing compared to 42% and <0.5% accounted for by out-of-pocket payments (again due to its regressivity) and private health insurance contributions respectively.

The results of the incidence analysis shown in Figure 6.2 do not provide a full picture of progressivity across the entire income distribution parade. Only five quintiles were categorised. Variations in incomes across these quintiles are not taken into account. Also results could be sensitive to how households located at the bounds of each quintile are handled or the number of quantiles used. In Figure 6.3 the Lorenz curve of income distribution and concentration curves of out-of-pocket payments, private health insurance and general taxes are shown. These curves also indicate that general taxes are progressive because its concentration curve lies below the Lorenz curve. Similarly private health insurance is generally progressive. Out-of-pocket payment is fairly proportional as its concentration curve almost coincides with the Lorenz curve.

Figure 6.3: Concentration curves of health care payments, South Africa



While the results presented in Figures 6.2 and 6.3 are informative, they are still descriptive in nature. This is because they do not indicate the precise extent of progressivity or regressivity or their statistical significance. In this regard progressivity indices are used to show the extent of progressivity in each health care financing mechanism. As noted above Kakwani indices were computed for all the financing mechanisms. In Table 6.1 the cumulative shares of health care payments via all direct taxes by income quintiles are presented. First the distribution of personal income tax is presented followed by corporate income tax and then total direct taxes, which is a combination of personal income and corporate income taxes.

Table 6.1: Cumulative shares of payments (direct taxes) by income quintile, South Africa 2005/2006

Equivalent household income quintile	Equivalent household income	Direct taxes		
		Personal income tax	Corporate income tax	Total direct taxes
poorest 20%	2.47%	0.02% [†]	0.92% [†]	0.45% [†]
(standard error)	(0.0339)	(0.0045)	(0.2541)	(0.0634)
poorest 40%	7.20%	0.20% [†]	3.04% [†]	1.54% [†]
	(0.0913)	(0.0175)	(0.8403)	(0.2159)
poorest 60%	15.09%	0.82% [†]	6.31% [†]	3.41% [†]
	(0.1788)	(0.0473)	(1.7138)	(0.4488)
poorest 80%	30.73%	5.73% [†]	14.28% [†]	9.76% [†]
	(0.3191)	(0.2407)	(3.9374)	(1.3211)
test of dominance				
- against 45° line		-	-	-
- against Lorenz curve		-	-	-
Concentration Index ^a	0.6377*	0.8833*	0.7969*	0.8426*
(standard error)	(0.0091)	(0.0098)	(0.0248)	(0.0106)
Kakwani Index		0.2457*	0.1593*	0.2050*
(standard error)		(0.0097)	(0.0275)	(0.0118)

Notes:

For shares: **bold** indicates significant difference from population share (5%)

[†] indicates significant difference from income share (5%).

* indicates statistical significance at 1% level.

Dominance tests:

- indicates the 45 degree line / Lorenz curve dominates the concentration curve

+ indicates concentration curve dominates 45 degree line / Lorenz curve

[] blank indicates non-dominance.

Dominance is rejected if there is at least one significant difference in one direction and no significant

difference in the other, with comparisons at 19 quantiles and 5% significance level.

a. Gini index for equivalent household income.

The overall extent of income distribution is presented under the column “equivalent household income.” This shows a highly unequal distribution of income where the poorest 20% of the population accounts for only about 2.5% of all income in South Africa while the top 20% of the population accounts for about 69%. The bottom 80% of the population only has about 31% of the total income in South Africa. Across all quintiles the shares of income are significantly different from their population shares. A summary measure of relative income inequality using Equation (6.1) is the Gini index. By definition a Gini index of zero indicates perfect equality in the distribution of income where the bottom $x\%$ of the population receives $x\%$ of all income; and a Gini index of one indicates the most unequal distribution where all

income is concentrated on one individual. Basically the closer the value of the Gini index to one, the more unequal is the distribution of income. In Table 6.1 the Gini index of income inequality was estimated at 0.64. This summary measure was statistically different from zero at 1% level of significance and reconfirms the distribution implied by the cumulative shares that income distribution is highly unequal in South Africa. This distribution of income will affect the extent of progressivity/regressivity of health care financing in South Africa.

Turning to the components of direct taxes as shown in Table 6.1, the distributions of health care financing via personal income tax and corporate income tax are progressive. The bottom 20% of the population that accounts for about 2.5% of total income contributes only about 0.02% of total personal income tax revenue and about 0.9% of corporate income tax revenue spent on health care in South Africa. These differences were estimated to be statistically significant and they show that the contributions by the bottom 20% of the population are less than their income share. Also the poorest 40%, 60% and 80% that respectively account for 7.2%, 15.1% and 30.7% of total income in the country contribute about 0.2%, 0.8% and 5.7% of the total personal income tax revenue and 3%, 6.3% and 14.3% of total corporate tax revenue spent on health care financing. The bulk of total personal income tax revenue (>94%) and corporate income tax revenue (>85%) used to finance health care is borne by the richest 20% of the population. The burden borne by the richest 20% of the population is however more than their share of total income. The distribution of personal income tax burden shown in Table 6.1 also reflects that of personal income tax payers in South Africa. Based on the distribution of personal income tax payers estimated from the IES only about 1% of personal income tax payers are in the bottom 20% of the population compared to about 59% in the top 20% of income distribution.

Combining personal income tax and corporate income tax shows that total direct taxes are also progressive. The cumulative shares of payments in Table 6.1 indicate that while the bottom 20% of the population accounts for 2.5% of total income, they contribute only 0.45% of total direct taxes used to finance health care in South Africa. Similarly the bottom 40% contributes 1.54% of total direct tax revenue but earn only 7.2% of total income. Again about 90% of total direct tax revenue comes from the top 20% of the population. This proportion for the richest quintile is greater than their share of total income. The difference between shares of health care payments via tax and income shares were statistically significant at 5% level. Overall these cumulative shares imply that financing health care via direct taxes is progressive in South Africa.

The cumulative distribution of health care financing burden for indirect taxes in Table 6.2 shows that fuel levy, value-added tax and excise taxes are regressive. The poorest 20% of the population that earn 2.5% of total income contributes significantly more than their income share in indirect taxes to finance health care. For instance they contribute 3.3%, 2.7% and 3.9% of total fuel levy, value-added tax and excise taxes revenue respectively that is used to finance health care. The cumulative shares of payments for all the indirect taxes at different quantiles are significantly greater than that for income. Of all these indirect taxes, value-added tax has the lowest cumulative shares across all quantiles signifying that it is the least regressive, while excise taxes are the most regressive.

Table 6.2: Cumulative shares of payments (indirect taxes) by income quintile, South Africa 2005/2006

Equivalent household income quintile	Equivalent household income	Indirect taxes			
		Fuel levy	Value-added tax (VAT)	Excise tax	Total indirect taxes
poorest 20%	2.47%	3.27% [†]	2.69% [†]	3.87% [†]	2.90% [†]
(standard error)	(0.0339)	(0.0580)	(0.0539)	(0.1793)	(0.0540)
poorest 40%	7.20%	9.01% [†]	8.21% [†]	11.73% [†]	8.71% [†]
	(0.0913)	(0.1328)	(0.1453)	(0.4258)	(0.1403)
poorest 60%	15.09%	18.52% [†]	17.74% [†]	26.10% [†]	18.78% [†]
	(0.1788)	(0.2434)	(0.3013)	(0.7996)	(0.2847)
poorest 80%	30.73%	37.54% [†]	35.34% [†]	52.94% [†]	37.60% [†]
	(0.3191)	(0.4233)	(0.5705)	(1.3814)	(0.5344)
test of dominance					
- against 45° line		-	-	-	-
- against Lorenz curve		+	+	+	+
Concentration Index ^a	0.6377*	0.5642*	0.6002*	0.4465*	0.5781*
(standard error)	(0.0091)	(0.0073)	(0.0070)	(0.0161)	(0.0067)
Kakwani Index		-0.0734*	-0.0374*	-0.1912*	-0.0595*
(standard error)		(0.0104)	0.0082	(0.0195)	(0.0087)

Notes:

For shares: **bold** indicates significant difference from population share (5%)

[†] indicates significant difference from income share (5%).

* indicates statistical significance at 1% level.

Dominance tests:

- indicates the 45 degree line / Lorenz curve dominates the concentration curve

+ indicates concentration curve dominates 45 degree line / Lorenz curve

[] blank indicates non-dominance.

Dominance is rejected if there is at least one significant difference in one direction and no significant

difference in the other, with comparisons at 19 quantiles and 5% significance level.

a. Gini index for equivalent household income.

Unlike direct taxes shown in Table 6.1, for indirect taxes the tax burden on the top 20% of the population is significantly less than their income share. About 62%, 65% and 47% respectively of total fuel levy, value-added tax revenue and excise taxes used in financing health care is accounted for by the top 20% of the population. The distributional pattern found in indirect taxes reflects the broad nature of these taxes. In general total indirect taxes that combine fuel levy, value-added tax and excise taxes show a regressive pattern. The cumulative shares of payments across the quantiles are significantly higher than those of the corresponding income distribution quantiles. The bottom 20% of the population accounts for about 3% of total indirect tax revenue used in financing health care compared to 2.5% of total income. The bottom 80% of the population contributes about 37.6% of total indirect taxes but accounts for about 30.7% of total income.

Table 6.3 presents the distribution of health care financing from general taxes (a combination of all indirect and direct taxes), out-of-pocket payments and private health insurance contributions. In Tables 6.1 and 6.2 we saw that the distribution of direct taxes is progressive while that of indirect taxes is regressive. The distribution of general taxes will therefore depend on these two distributions and their relative shares of general taxes. As shown in Table 6.3 apart from out-of-pocket payments, the distributions of general taxes and private health insurance are progressive. About 1.4% of general tax revenue spent on health care in South Africa comes from the bottom 20% of the population compared to 79.8% coming from the top 20% of the population. The bottom 40%, 60% and 80% respectively account for about 4%, 9% and 20% of general tax revenue spent on health care financing. The share for each quantile is significantly less than the corresponding income share. The progressive nature of general taxes can be attributed to the greater progressivity of direct taxes compared to the regressivity of indirect taxes and the fact that direct taxes contribute 57% of general taxes while indirect taxes contribute 43%.

Table 6.3: Cumulative shares of health care payments by income quintile, South Africa 2005/2006

Equivalent household income quintile	Equivalent household income	General taxes	Out-of-pocket payments	Private health Insurance	Overall health care finance
poorest 20% (standard error)	2.47% (0.0339)	1.37% [†] (0.1142)	2.79% (0.1855)	0.01% [†] (0.0024)	0.94% [†] (0.0377)
poorest 40%	7.20% (0.0913)	4.24% [†] (0.3509)	7.71% (0.4861)	0.07% [†] (0.0147)	2.80% [†] (0.1072)
poorest 60%	15.09% (0.1788)	9.20% [†] (0.7524)	16.55% (1.0460)	0.92% [†] (0.0970)	6.38% [†] (0.2431)
poorest 80%	30.73% (0.3191)	20.25% [†] (1.6599)	29.80% (1.8341)	10.07% [†] (0.4915)	16.86% [†] (0.6457)
test of dominance					
- against 45° line		-	-	-	-
- against Lorenz curve		-	[]	-	-
Concentration Index ^a (standard error)	0.6377* (0.0091)	0.7431* (0.0123)	0.6361* (0.0198)	0.8069* (0.0117)	0.7576* (0.0080)
Kakwani Index (standard error)		0.1054* (0.0115)	-0.0015 (0.0191)	0.1692* (0.0163)	0.1199* (0.0108)

Notes:

For shares: **bold** indicates significant difference from population share (5%)

[†] indicates significant difference from income share (5%).

* indicates statistical significance at 1% level.

Dominance tests:

- indicates the 45 degree line / Lorenz curve dominates the concentration curve

+ indicates concentration curve dominates 45 degree line / Lorenz curve

[] blank indicates non-dominance.

Dominance is rejected if there is at least one significant difference in one direction and no significant difference in the other, with comparisons at 19 quantiles and 5% significance level.

a. Gini index for equivalent household income.

The distribution of out-of-pocket health care expenditures on the other hand is not clear-cut and the difference between the cumulative shares of out-of-pocket payments and income are not statistically significant across all quantiles. For example the bottom 20% of the population contributes about 2.8% of total out-of-pocket expenditures but they account for about 2.5% of total income signifying that their payment share is greater than their income share. The same pattern is true for the bottom 40% and 60% of the population. However for the bottom 80%, the out-of-pocket payments share is slightly lower than the income share.

The distribution of private health insurance like general taxes shows a very progressive pattern. About 0.01% of total private insurance contributions come from the bottom 20% of the population while 90% of this contribution comes from the top 20% of the population. The cumulative share of private health insurance contributions at each quantile is significantly lower than that for the corresponding quantile of income distribution. As noted earlier because the

distribution of the insured population is skewed towards the top quintiles, the contribution pattern reflects this underlying distribution of insured individuals. Only less than 1% of total private health insurance contributions come from the poorest 60% of the population.

Overall as shown in Table 6.3 the distribution of total health care financing in South Africa is progressive. This is because the cumulative share of total health care payments for the bottom $x\%$ of the population is significantly less than their cumulative share of their income at the 5% level. For instance the bottom 20% (80%) of the population has a total health care payment share of about 1% (17%) compared to total income share of 2.5% (31%).

6.4 FORMAL ASSESSMENT OF THE PROGRESSIVITY OF HEALTH CARE FINANCING IN SOUTH AFRICA

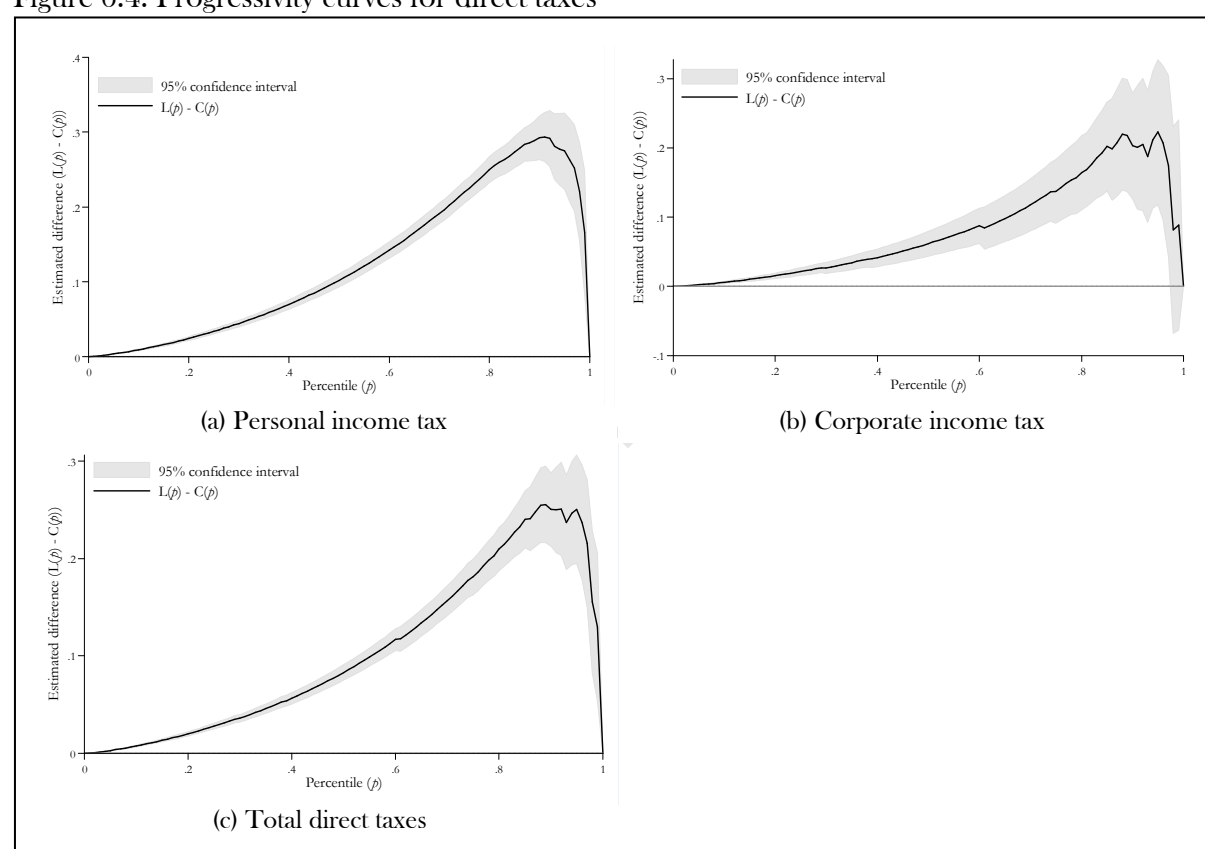
The preliminary results based on Tables 6.1 – 6.3 show that indirect taxes are regressive, direct taxes are progressive, general taxes are progressive, private health insurance contributions are progressive, out-of-pocket payments are roughly proportional and overall health care financing is progressive. The same conclusions can be drawn from the graphs in Figures 6.2 and 6.3. However we need to provide statistical indices that summarise the extent of progressivity for each of these financing mechanisms. For example we want to know the extent to which private health insurance, general taxes and overall health care financing are progressive. We also want to know if the progressive relationship is statistically significant. Using the concentration indices of payments and Gini index of income inequality, the extent of progressivity is assessed and presented in Tables 6.1 – 6.3. Tests of dominance between the concentration curve and the line of equality (i.e. the 45° line) and between the concentration curve and the Lorenz curve are also presented.

6.4.1 Direct taxes

The progressivity of direct taxes as noted earlier results from the progressivity of personal income tax and corporate income tax. The concentration index of personal income tax was estimated at 0.88 and this is greater than the Gini index of 0.64. This gives a positive and statistically significant Kakwani index of 0.25. The significant Kakwani index confirms that personal income tax as a source of health care financing in South Africa is progressive. The stochastic dominance test of the concentration curve and the Lorenz curve also confirms that

the distribution of personal income tax (i.e. via the concentration curve) is everywhere ($p \in [0,1]$) below the Lorenz curve of income distribution. However the magnitude of this positive difference may vary across the entire population distribution. To examine how this magnitude varies across the entire population distribution Figure 6.4(a) plots the progressivity curve for personal income tax. This Figure shows that while the difference ($L(p) - C_T(p)$) is everywhere positive, signifying a progressive distribution, progressivity is significantly larger at higher percentiles than at the lower percentiles. The confidence intervals around the estimates show a statistically significant difference.

Figure 6.4: Progressivity curves for direct taxes



Similarly for corporate income tax as shown in Table 6.1 the estimated concentration index is 0.797 and this is greater than the estimated Gini index of 0.64. The result is a positively significant Kakwani index of 0.16. The positive Kakwani index was confirmed by stochastic tests of dominance which shows that the concentration curve of corporate income tax is everywhere dominated by the Lorenz curve. The progressivity curve of corporate income tax as shown in Figure 6.4(b) indicates that though the Kakwani index is positive and the concentration curve is dominated by the Lorenz curve, the difference is larger at the top percentiles than at lower percentiles. The statistically significant confidence intervals around

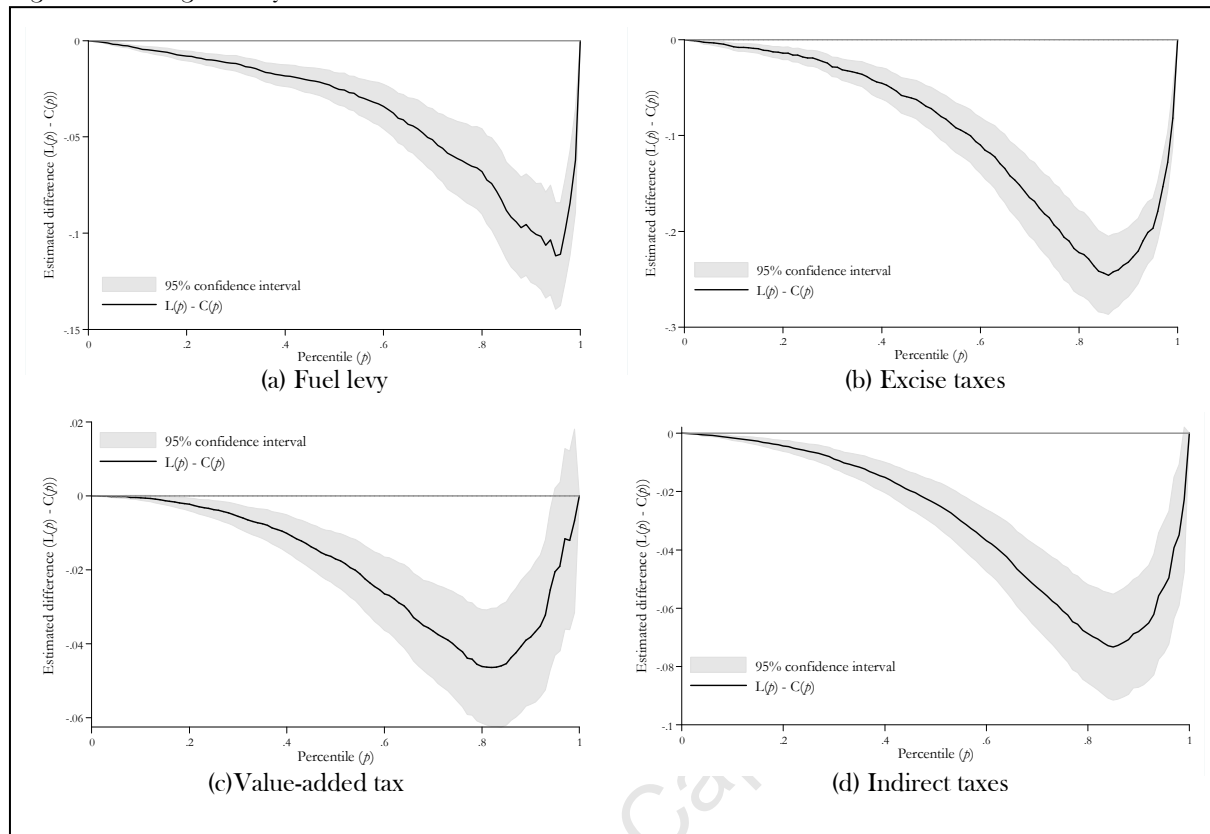
the estimated difference also confirm the results of the stochastic dominance that corporate income tax is everywhere progressive.

Following from these, it is not surprising therefore that financing health care via direct taxes is progressive. The estimated concentration index for direct taxes is 0.84 and this is again greater than the Gini index estimated at 0.64. The statistically significant Kakwani index of 0.2 estimated based on these indices confirms that general taxes are a progressive source of health care financing in South Africa. The extent of progressivity of total direct taxes is less than that of personal income tax but greater than that of corporate income tax. The stochastic dominance result also indicates that the Lorenz curve everywhere dominates the concentration curve of total direct taxes. As also graphically shown in Figure 6.4(c), the progressivity curve of direct taxes is everywhere positive. The confidence interval around the estimated difference is everywhere statistically significant and reconfirms the results of stochastic dominance. Progressivity was greater at higher percentiles than at lower percentiles.

6.4.2 Indirect taxes

Turning to indirect taxes, we see a clearly different pattern from that observed in direct taxes. Here the progressivity curves shown in Figure 6.5 are generally below the line of proportionality (i.e. the zero line). From the results in Table 6.2, the concentration index of fuel levy was estimated at 0.56. With the Gini index estimated at 0.64, the resulting Kakwani index is -0.07. The negative Kakwani index was observed to be significantly different from zero implying that fuel levy is a regressive source of health care financing in South Africa. Stochastic test of dominance also concluded that fuel levy is everywhere regressive and places a greater burden on the poor compared to the rich. As also shown in Figure 6.5(a) the progressivity curve of fuel levy is everywhere negative and statistically different from zero. The significant differences were larger in absolute values within the top percentiles than at lower percentiles which reflect the underlying inequality in income distribution in the country.

Figure 6.5: Progressivity curves for indirect taxes



Excise taxes that comprise alcohol taxes and tobacco taxes, have an estimated concentration index of 0.45 (see Table 6.2). This is statistically significant and lower than the Gini index of 0.64. The resulting Kakwani index is -0.19 and it is significantly different from zero. This index confirms that excise taxes are regressive sources of health care financing in South Africa. The regressivity of excise taxes is also confirmed by the stochastic test of dominance and the progressivity curve shown in Figure 6.5(b). Based on this Figure, excise taxes are seen to be significantly regressive across the entire population distribution. The regressivity implied in Figure 6.5(b) indicates that while excise taxes are significantly regressive across the entire distribution, regressivity is higher at the top part of the distribution compared to the lower parts. This is similar to the relationship observed for fuel levy.

The same pattern of regressivity observed in fuel levy and excise taxes are observed for value-added tax. The value of the concentration index (0.60) is less than the Gini index estimated at 0.64. A statistically significant value of the Kakwani index was estimated at -0.04. The negative value of Kakwani index, which signifies regressivity, was also confirmed by the test of dominance and the progressivity curve in Figure 6.5(c). The progressivity curve shows that the

distribution of value-added tax is everywhere regressive. The confidence interval bands around the progressivity curve indicate that the regressivity in value-added tax is everywhere statistically significant except for the top 5% of the population where the confidence interval bands cross the zero line. Again regressivity is higher at the top of the income distribution compared to the rest.

Because all the indirect taxes turned out to be regressive sources of financing health care in South Africa, their combination is no different. From Table 6.2 the concentration index of total indirect taxes (0.58), though positive, was less than the Gini index of income distribution. The difference between these indices produced a negative and statistically significant Kakwani index of -0.06. This confirms that indirect taxes are regressive in South Africa. The test of stochastic dominance also reconfirms that indirect taxes are regressive because the Lorenz curve is dominated by the concentration curve of indirect taxes. This dominance as also shown in Figure 6.5(d) is consistent and statistically significant across the entire distribution. Regressivity generally increases as you move to higher percentiles.

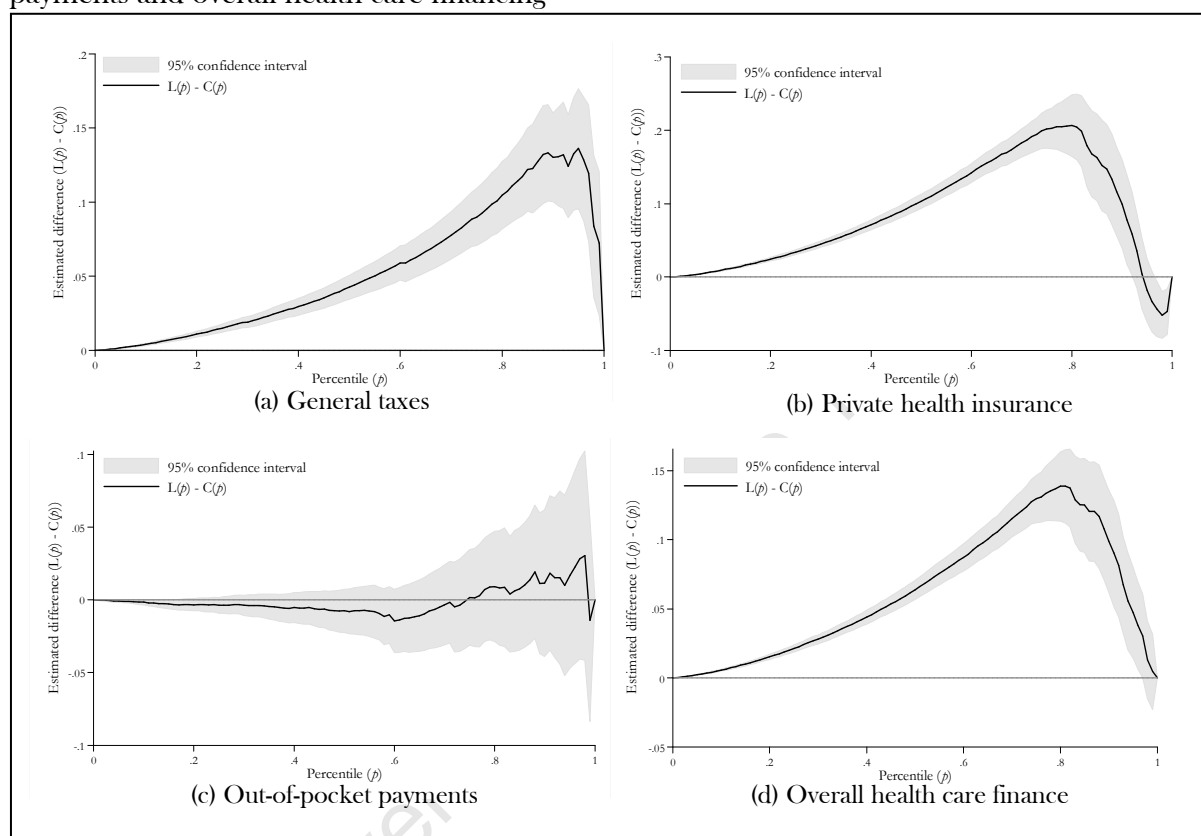
6.4.3 Overall health care financing

In South Africa, like in many other countries, because health care is financed from different sources, to understand the incidence of health care financing it is important to consider the progressivity of all the constituent parts which in this case include general taxes, out-of-pocket expenditures and private health insurance contributions.

As shown in Table 6.3 the concentration index for general taxes is estimated at 0.74 which is significantly greater than the Gini index of 0.64. The difference between these estimates gives a statistically significant Kakwani index of 0.105. This confirms that general taxes are progressive health care financing sources. The progressivity of general taxes is driven mainly by (i) the progressivity of direct taxes that more than offset the regressivity in indirect taxes and (ii) the relative share of direct taxes in total tax revenue. This conclusion is drawn because the Kakwani index is additively separable (Kakwani, 1977). The Kakwani index of direct taxes was estimated at 0.205 while that of indirect taxes was -0.060. Because direct taxes comprise over 60% of total tax revenue used in financing health care, and it is significantly progressive, the combined effect of the progressive direct tax and its larger share in general taxes resulted in a positive and statistically significant Kakwani index for general taxes. Test of dominance also established that the concentration curve of general taxes is everywhere dominated by the

Lorenz curve of income distribution. The progressivity curve shown in Figure 6.6(a) also indicates that the difference is everywhere positive and significantly different from zero. This difference that peaks at about 0.14 was estimated to be higher at top percentiles compared to lower percentiles.

Figure 6.6: Progressivity curves for general taxes, private health insurance, out-of-pocket payments and overall health care financing



Out-of-pocket expenditures that make up less than 14% of total health care financing in South Africa as shown in Table 6.3 has a positive and statistically significant concentration index of 0.636 that is slightly less than the Gini index estimated at 0.6377. The resultant Kakwani index is -0.0015. Though this indicates that out-of-pocket payments are regressive, this index is statistically not different from zero. In fact in absolute terms the standard error exceeds the value of the estimated index. This means that out-of-pocket payments are, statistically speaking, a proportional health care financing mechanism. This proportionality in out-of-pocket payments is confirmed by the stochastic dominance result which indicates lack of dominance between the two curves. The progressivity curve shown in Figure 6.6(c) indicates clearly that a regressive pattern was observed in the distribution of out-of-pocket payments up until the 75th percentile when it becomes progressive. However based on the 95% confidence interval bands, these differences were never statistically different from zero. The overall

conclusion based on the results therefore is that out-of-pocket payments are a proportional financing mechanism in South Africa.

Private health insurance like general taxes has a concentration index value that is greater than the Gini index. The concentration index (0.81) as shown in Table 6.3 is significantly different from zero and so is the resultant Kakwani index (0.17). This means that private health insurance is a progressive health care financing mechanism and places greater burden on the rich compared to the poor. The MCA to stochastic dominance indicates that the concentration curve of private health insurance is dominated by the Lorenz curve. However the visual inspection of the progressivity curve shown in Figure 6.6(b) indicates that the dominance is not entirely true for the top 5% of the distribution. What this means is that though private health insurance is progressive for those at the lower income distribution, it is regressive for the top 5% of income distribution. In fact it is significantly regressive for the top 2%. The multiple comparison approach truncates this part of the distribution in arriving at the dominance conclusion. As noted earlier only the distribution between the interval [0.05, 0.95] was used in testing the dominance based on the MCA approach. The progressivity curve also indicates that the difference ($L(p) - C_T(p)$) is higher at top percentiles and this is attributable to the skewed distribution of the insured.

Overall as shown in Table 6.3 the concentration index of health care financing in South Africa is positive and statistically different from zero. The Kakwani index is estimated at 0.12 and it is also statistically significant at 1% level. Because overall health care financing is a weighted combination of general taxes, out-of-pocket payments and private health insurance, the progressive pattern is expected. This is because all other sources apart from out-of-pocket payment, which is proportional, are progressive. The relative progressivity of all these sources and their relative contributions to total health care financing determine the overall extent of progressivity. Stochastic test of dominance result also indicates that the Lorenz curve dominates the concentration curve of total health care financing. The progressivity curve in Figure 6.6(d) shows that the Lorenz curve is everywhere significantly above the concentration curve of total health care financing in South Africa. The progressivity relationship is significantly higher at the top of the income distribution than at lower levels.

The result presented in Table 6.3 for total health care financing has not accounted for a small proportion of general taxes that was not estimated. As noted earlier in Chapter 5 this makes up about 3.7% of total health care financing and includes, among others, taxes on property and

unidentified levies, stamp duties and fines, air departure tax and skills development levy. In order to arrive at an index for overall health care financing that takes this into account, three assumptions were used to allocate this proportion based on the estimated Kakwani indices for the other financing sources. The first case assumes that the proportion that was not estimated is distributed as the weighted average of all the allocated taxes; the second assumes that it is distributed as the weighted average of value-added tax; and the third as an excise tax. The additively separable property of the Kakwani index makes it easy to compute the adjusted Kakwani index in Table 6.4 as the weighted sum of all the other Kakwani indices using the revised shares (i.e. macro weights) as the weights for each case scenario.

Basically the adjusted Kakwani index π_K^{adj} was obtained using the estimated Kakwani index of each health care financing source (π_j) and the corresponding revised weights (w_j) in Table 6.4 as:

$$\pi_K^{\text{adj}} = \sum_j w_j \pi_j$$

As shown in Table 6.4 the overall Kakwani index is robust to these assumptions. The changes to the adjusted Kakwani indices are very minimal. For example case 1 produced an overall adjusted Kakwani index of 0.118 which is comparable to that shown in Table 6.3. The smallest value of Kakwani index was however recorded using the third assumption.

Table 6.4: Health care financing by source and Kakwani for total payments, South Africa 2005/2006

Finance source	Share of total finance	Method of allocation	Kakwani index by source	Macro weights		
				Case 1	Case 2	Case 3
General taxes	40.41 %					
a. Personal income tax	0.121	estimated	0.246	0.134	0.121	0.121
b. Value-added tax	0.109	estimated	-0.037	0.120	0.147	0.109
c. Corporate income tax	0.101	estimated	0.159	0.111	0.101	0.101
d. Fuel levy	0.020	estimated	-0.073	0.022	0.020	0.020
e. Excise tax	0.015	estimated	-0.191	0.017	0.015	0.053
f. Other	0.037	allocated	-			
Private health insurance	45.27 %	reported	0.169	0.453	0.453	0.453
Out-of-pocket payments	14.32 %	reported	-0.002	0.143	0.143	0.143
Total	100.00 %			1.00	1.00	1.00
% general taxes fully estimated	96.26%					
Adjusted Kakwani index for total health care finance				0.1176	0.1124	0.1066

Adjustment scenarios for macro weights:

Case 1 - Unallocated tax (3.74%) distributed as the weighted average of all allocated taxes.

Case 2 - Unallocated tax (3.74%) distributed as value-added tax.

Case 3 - Unallocated tax (3.74%) distributed as excise tax.

Kakwani indices presented in Tables 6.1 – 6.3 cannot be used to order relative progressivity of these financing mechanisms. Formal tests of dominance are required in order to determine which of these financing mechanisms is statistically more progressive or less regressive than the other. These conclusions are important from a policy perspective in order to reform the health care financing system. Generally the financing mechanisms that are *more progressive* and *equitable* are promoted over those that are *less progressive*, *regressive* or *inequitable*. The results of the dominance test using the MCA are presented in Table 6.5. As shown excise taxes are the most regressive of all health care financing mechanisms because it has the least value of Kakwani index (-0.19) and its concentration curve is everywhere above those of other health care financing mechanisms. Fuel levy is the next most regressive source of health care financing. Its concentration curve dominates those of other financing sources except for that of excise taxes. Value-added tax and out-of-pocket payments are the third most regressive financing sources. This is because their concentration curves coincide. The most progressive health care financing sources are private health insurance and personal income tax. Even though their concentration curves cross each other, they are dominated by those of other sources. The Kakwani index of personal income tax as shown in Table 6.1 is however higher than that of private health insurance shown in Table 6.3 but we cannot statistically conclude which among these two is more progressive. Also between out-of-pocket payments, private health insurance and general taxes as shown in Table 6.5, the concentration curve of general taxes is dominated by that of out-of-pocket payments but it crosses that of private health insurance. These findings imply that general taxes are more progressive than out-of-pocket payments but we cannot statistically assess its relative progressivity with private health insurance because private health insurance is significantly regressive at the top decile.

Table 6.5: Results of dominance tests, South Africa 2005/2006

	Fuel levy	Value added tax	Out-of-pocket payments	Corporate income tax	Private health insurance	Personal income tax	General taxes
Excise tax	D	D	D	D	D	D	-
Fuel levy		D	D	D	D	D	-
Value-added tax			non-D	D	D	D	-
Out-of-pocket payments				D	D	D	D
Corporate income tax					non-D	non-D	-
Private health insurance						CC	CC

Notes:

D indicates that concentration curve of row source dominates (is more progressive than) that of column source.

Dominance is rejected if there is at least one significant difference in one direction and no significant difference in the other, with comparisons at 19 quantiles and 5% level of significance.

non-D indicates that non-dominance between the concentration curves cannot be rejected.

CC indicates that both concentration curves cross.

A summary of the assessment of progressivity is contained in Table 6.6.

Table 6.6: Cumulative shares of health care payments by income quintile, South Africa 2005/2006

Equivalent Household Income Quintile	Equivalent Household Income	Direct taxes			Indirect taxes				General taxes	OOP payments	Private health Insurance	Overall health care finance
		Personal income Tax	Corporate income tax	Total direct taxes	Fuel levy	Value- added-tax	Excise tax	Total indirect taxes				
poorest 20%	2.47%	0.02% [†]	0.92% [†]	0.45% [†]	3.27% [†]	2.69% [†]	3.87% [†]	2.90% [†]	1.37% [†]	2.79%	0.01% [†]	0.94% [†]
(standard error)	(0.0339)	(0.0045)	(0.2541)	(0.0634)	(0.0580)	(0.0539)	(0.1793)	(0.0540)	(0.1142)	(0.1855)	(0.0024)	(0.0377)
poorest 40%	7.20%	0.20% [†]	3.04% [†]	1.54% [†]	9.01% [†]	8.21% [†]	11.73% [†]	8.71% [†]	4.24% [†]	7.71%	0.07% [†]	2.80% [†]
	(0.0913)	(0.0175)	(0.8403)	(0.2159)	(0.1328)	(0.1453)	(0.4258)	(0.1403)	(0.3509)	(0.4861)	(0.0147)	(0.1072)
poorest 60%	15.09%	0.82% [†]	6.31% [†]	3.41% [†]	18.52% [†]	17.74% [†]	26.10% [†]	18.78% [†]	9.20% [†]	16.55%	0.92% [†]	6.38% [†]
	(0.1788)	(0.0473)	(1.7138)	(0.4488)	(0.2434)	(0.3013)	(0.7996)	(0.2847)	(0.7524)	(1.0460)	(0.0970)	(0.2431)
poorest 80%	30.73%	5.73% [†]	14.28% [†]	9.76% [†]	37.54% [†]	35.34% [†]	52.94% [†]	37.60% [†]	20.25% [†]	29.80%	10.07% [†]	16.86% [†]
	(0.3191)	(0.2407)	(3.9374)	(1.3211)	(0.4233)	(0.5705)	(1.3814)	(0.5344)	(1.6599)	(1.8341)	(0.4915)	(0.6457)
test of dominance												
- against 45° line		-	-	-	-	-	-	-	-	-	-	-
- against Lorenz curve		-	-	-	+	+	+	+	-	[]	-	-
Concentration Index ^a	0.6377*	0.8833*	0.7969*	0.8426*	0.5642*	0.6002*	0.4465*	0.5781*	0.7431*	0.6361*	0.8069*	0.7576*
(standard error)	(0.0091)	(0.0098)	(0.0248)	(0.0106)	(0.0073)	(0.0070)	(0.0161)	(0.0067)	(0.0123)	(0.0198)	(0.0117)	(0.0080)
Kakwani Index		0.2457*	0.1593*	0.2050*	-0.0734*	-0.0374*	-0.1912*	-0.0595*	0.1054*	-0.0015	0.1692*	0.1199*
(standard error)		(0.0097)	(0.0275)	(0.0118)	(0.0104)	0.0082	(0.0195)	(0.0087)	(0.0115)	(0.0191)	(0.0163)	(0.0108)

Notes:

For shares: **bold** indicates significant difference from population share (5%)

[†] indicates significant difference from income share (5%).

* indicates statistical significance at 1% level.

Dominance tests: - indicates the 45 degree line / Lorenz curve dominates the concentration curve

+ indicates concentration curve dominates 45 degree line / Lorenz curve

[] blank indicates non-dominance.

Dominance is rejected if there is at least one significant difference in one direction and no significant difference in the other, with comparisons at 19 quantiles and 5% significance level.

a. Gini index for equivalent household income.

6.5 SENSITIVITY AND ROBUSTNESS ANALYSES OF THE INCIDENCE RESULTS

In this section, some robustness checks and sensitivity analyses were performed on the results presented. This basically requires that certain assumptions made are subjected to rigorous scrutiny so as to ensure that the results obtained are robust to any sensible changes made to these assumptions. For example the results presented are based on certain assumptions about economies of scale and the consumption of a child relative to an adult. Also there is an underlying value judgement in the construction of the concentration and Gini indices used to assess the incidence of health care financing. The main question that is answered in this section is: Are the results still robust to changes in these assumptions within reasonable limits such that conclusions drawn remain valid?

First, sensitivity analysis was conducted on the Kakwani index by varying society/policy maker's aversion to inequalities, and then the robustness of the results was assessed to changes in assumptions on the equivalence scale parameters. It also assesses the robustness of the overall Kakwani index of progressivity to the assumptions about the economic incidence of corporate income tax.

6.5.1 Robustness of the Kakwani index to changes in aversion to inequality parameter

The formulation of the Kakwani index used in the analysis and presentation of the results in Tables 6.1 – 6.3 is based on an underlying ethical/value judgement. The basic assumption is that deviations from proportionality (i.e. $[L(p_i) - C_T(p_i)]$) are weighted equally across all units irrespective of the individual's location on the income distribution parade (i.e. the individual's percentile or rank). This means that all deviations are 'equally' important (Duclos and Araar, 2006).

If we write down a formulation for the traditional or standard Gini index as:

$$G = \int_0^1 2(p - L(p)) dp = 2 \int_0^1 (p - L(p)) dp \quad (6.5)$$

it is clear that the integral (or sum in a discrete case) weights all deviations by a constant multiplicative scalar of 2. This implicit value judgement may not be appealing to all because of its egalitarian orientation. We may be interested in a function that weights the deviations $(p - L(p))$ differently depending on the value of p in such a way that it incorporates some notion of

vertical equity. This means that the poor are prioritised in aggregating inequalities. One can then think of applying generalised percentile-dependent or rank-dependent weights that would be flexible to accommodate different ethical or value judgements (including egalitarian judgements) of decision makers. These weights would be used to weight the deviations ($p - L(p)$) and should be dependent on the location of the individual on the income distribution parade. One such class of weights that has been used in the inequality literature is the single parameter ethical weight (Donaldson and Weymark, 1980, 1983) that is dependent on one ‘ethical’ parameter (ν). The weight can also be viewed as the policy maker’s or society’s level of aversion or concern for inequalities of the poor relative to those of the rich (Yaari, 1988). A common specification is given as:

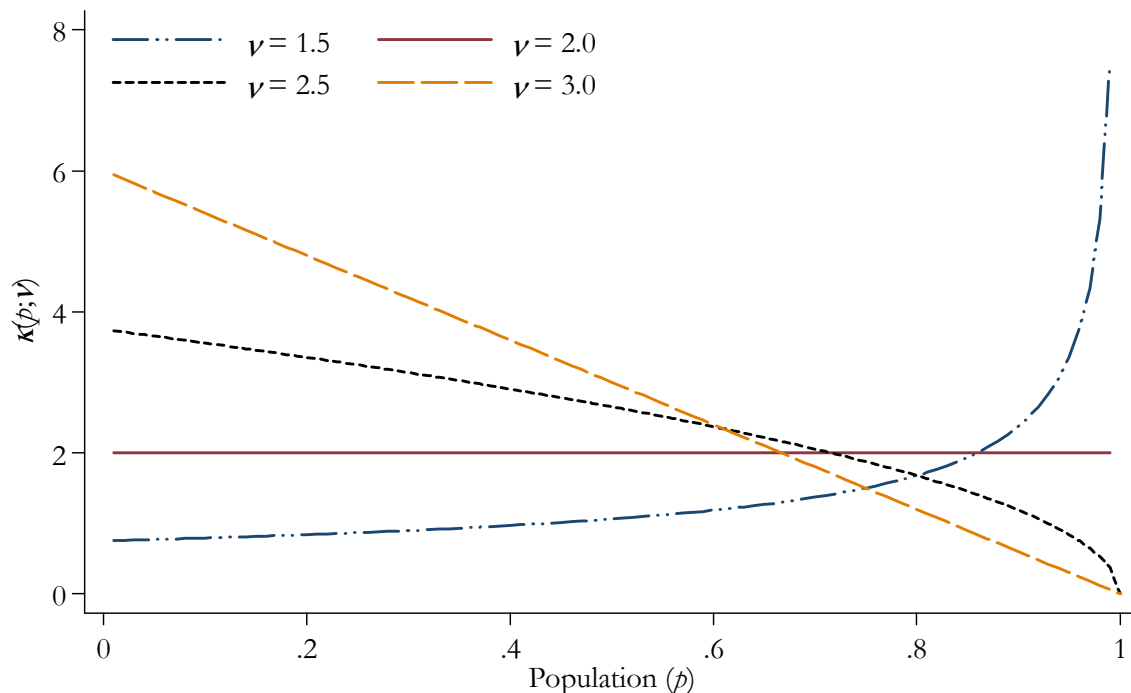
$$\kappa(p; \nu) = \nu(\nu - 1)(1 - p)^{(\nu - 2)} \quad (6.6)$$

where $\kappa(p; \nu)$ is everywhere positive for values of $\nu > 1$.

The need for such assessment is derived from the ethical judgements that underpin indices of relative inequality. Blackorby and Donaldson (1978) noted that since there are various value judgements underlying indices of relative inequality, it is important therefore to analyse the distributive impact of health care financing using different ‘ethical’ viewpoints. Further, Yaari (1988 pp.391-392) notes that the policy maker or society’s “concern for (in)equality is greater when the population is predominantly poor than when it is predominantly rich” as is the case in South Africa.

The weight implied by Equation (6.6) is illustrated in Figure 6.7 for values of $\nu = 1.5, 2.0, 2.5$ and 3.0 . The figure shows that higher values of $\nu > 2$ accord larger weight to the population at lower ranks than those at the top while lower values of $\nu < 2$ give lesser weights to the population at lower ranks relative to those at higher ranks. In general increasing the value of ν increases the weights assigned to the population at lower ranks and concurrently decreases those assigned to the population at higher ranks. If $\nu = 2$ the weight assigned to everyone in the population is the same irrespective of their rank. In the case when the value of $\nu > 2$, the policy maker or the society expresses more concern for the inequality of the poor. This is because in aggregating inequality into an index, higher weights are attached to deviations occurring at lower percentiles than those at higher percentiles. A similar weighting function has recently been applied to the assessment of catastrophic health care expenditures (see Ataguba, 2012b).

Figure 6.7: Single parameter ethical weighting function ($\kappa(p; \nu)$)



Applying the weighting function (6.6) to the standard Gini index in (6.5) gives:

$$G(\nu) = \int_0^1 (p - L(p)) \kappa(p; \nu) dp \quad (6.7)$$

where all members remain as previously defined. This is usually referred to in the literature as the \mathcal{S} -Gini index (i.e. single-parameter Gini index) (Duclos and Araar, 2006) or a variant of the extended Gini by Yitzhaki (1983). If $\nu = 2$ and we apply (6.6) we observe that $\kappa(p; \nu) = 2 \quad \forall p$ and (6.7) is equivalent to (6.5).

Analogously, we can define the \mathcal{S} -Gini type concentration index $C_T(\nu)$ as:

$$C_T(\nu) = \int_0^1 (p - C_T(p)) \kappa(p; \nu) dp \quad (6.8)$$

In relation to assessing the incidence of health care financing in South Africa we can write the \mathcal{S} -Gini type Kakwani index following (6.3) as:

$$\pi_K(\nu) = (L(p; \nu) - C_T(p; \nu)) \quad (6.9)$$

or equivalently as:

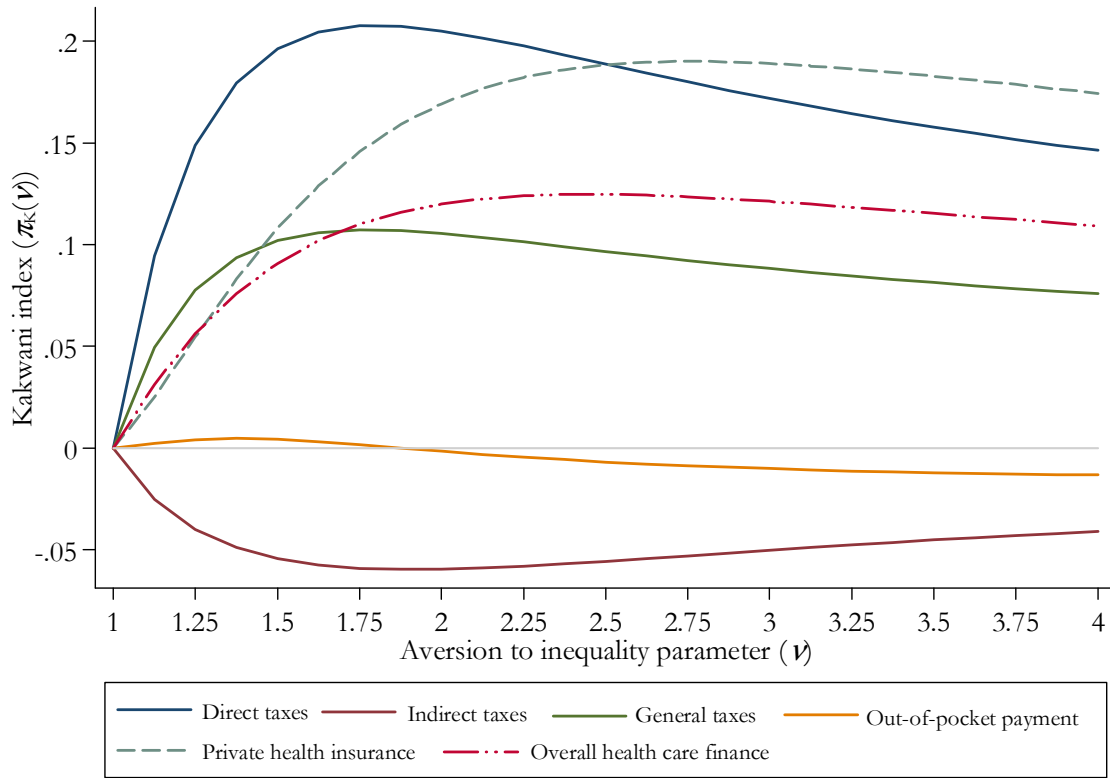
$$\pi_K(\nu) = C_T(\nu) - G(\nu) = (\hat{\xi}(\nu) / \hat{\mu}) - (\hat{\xi}_T(\nu) / \hat{\mu}_T) \quad (6.10)$$

Positive values of $\pi_K(\nu)$ indicate progressivity while negative values indicate regressivity. The standard Kakwani indices presented earlier simply weights all deviations from proportionality equally based on the value judgement that $\nu = 2$. This means that when $\nu = 2$, Equation (6.10) can simply be written as:

$$\pi_K(\nu = 2) = C_T - G = (\hat{\xi} / \hat{\mu}) - (\hat{\xi}_T / \hat{\mu}_T) \quad (6.11)$$

The impact of using different value judgements in the assessment of progressivity in health care financing in South Africa is presented in Figure 6.8. Using the general formulation of the Kakwani index in Equation (6.3), as shown in Figure 6.8, irrespective of the choice of parameter of inequality aversion direct taxes, general taxes, private health insurance and overall health care financing are progressive. This means that the observed progressivity of these financing sources is robust to the choice of the ethical parameter or value judgement used in aggregating deviations from proportionality. For values of the parameter $\nu < 2$ the value of the Kakwani indices were increasing with the value of the parameter. This is especially the case for private health insurance and total health care financing. It is important to note that as shown in Figure 6.7 when $\nu < 2$, we have that $\kappa(p; \nu) > 0 \forall p$ and $\partial \kappa(p; \nu) / \partial p > 0$. This means that deviations at the top part of the distribution are weighted more than those at the lower part. However when $\nu > 2$, we have that $\kappa(p; \nu) > 0 \forall p$ and $\partial \kappa(p; \nu) / \partial p < 0$ meaning that deviations at the lower part of the distribution receive more weight than those at the top. At values of $\nu < 2.5$ we notice that the curve for direct taxes dominates all other curves. However at $\nu = 2.5$ it crosses the curve for private health insurance. Thereafter private health insurance becomes ‘more progressive’ than direct taxes. Similarly when $\nu < 1.7$ the curve for general taxes lies above that of overall health care financing but this relationship switched thereafter and overall health care financing becomes ‘more progressive’ than general taxes.

Figure 6.8: The impact on progressivity of changes in the aversion to inequality parameter



Indirect taxes were seen to be everywhere regressive (i.e. $\pi_K(\nu) < 0$) irrespective of the choice of ν . However out-of-pocket expenditures were mildly progressive for values of $\nu < 2$ but they become regressive thereafter. As seen earlier in Figure 6.6(c) out-of-pocket payments are statistically speaking proportional. Also $L(p) - C_T(p) < 0$ at lower values of p but positive at higher values. Therefore if $\nu < 2$ it weights the positive deviations more than the negatives and that would result in a positive Kakwani index. When $\nu > 2$ the weight attached to the regressivity at lower percentiles more than offset the weight attached to the progressivity at higher percentiles resulting in overall negative Kakwani indices for out-of-pocket expenditures.

As shown in Figure 6.8 the underlying value judgements impact not so much on the general conclusions of progressivity or regressivity but on the ordering of the results. Overall, total health care financing is less responsive and therefore robust to the choice ν when $\nu \geq 2$. It is important to note here that for $\nu < 2$ the society weights inequality of the rich more than that of the poor and this is unlikely to be a popular consensus given the current level of income inequality and poverty in the country. This also confirms Yaari's (1988) assertion. Therefore

values of $\nu > 2$ would be plausible and the results based on these values, especially for values of $\nu > 2.5$ are consistent.

6.5.2 Robustness of the Kakwani index to changes in equivalence scale parameters

As noted earlier, results of inequality could be sensitive to the choice of equivalent adult scale parameters (Coulter *et al.*, 1992, Jenkins and Cowell, 1994, Aaberge and Melby, 1998). Therefore sensitivity analyses are important (Aaberge and Melby, 1998). It has been recognised that “[t]esting the sensitivity of inequality and poverty results to changes in the incorporation of needs would seem particularly important for those comparisons whose results can influence redistributive policies” (Duclos and Araar, 2006 p.31). Therefore important questions to ask include: what is the impact of changing economies of scale on the assessment of progressivity of health care financing in South Africa? Does it mean that progressivity is improved when families and households can live more cheaply? What difference does it make when we change our assumptions about the cost of children in intra-household resource allocation in South Africa? Does progressivity improve when we weight children far lower than adults? What is the combined effect on progressivity of concurrently changing both the assumptions about the economies of scale and the cost of children? Are the results robust to these assumptions and are they similar across the various financing mechanisms?

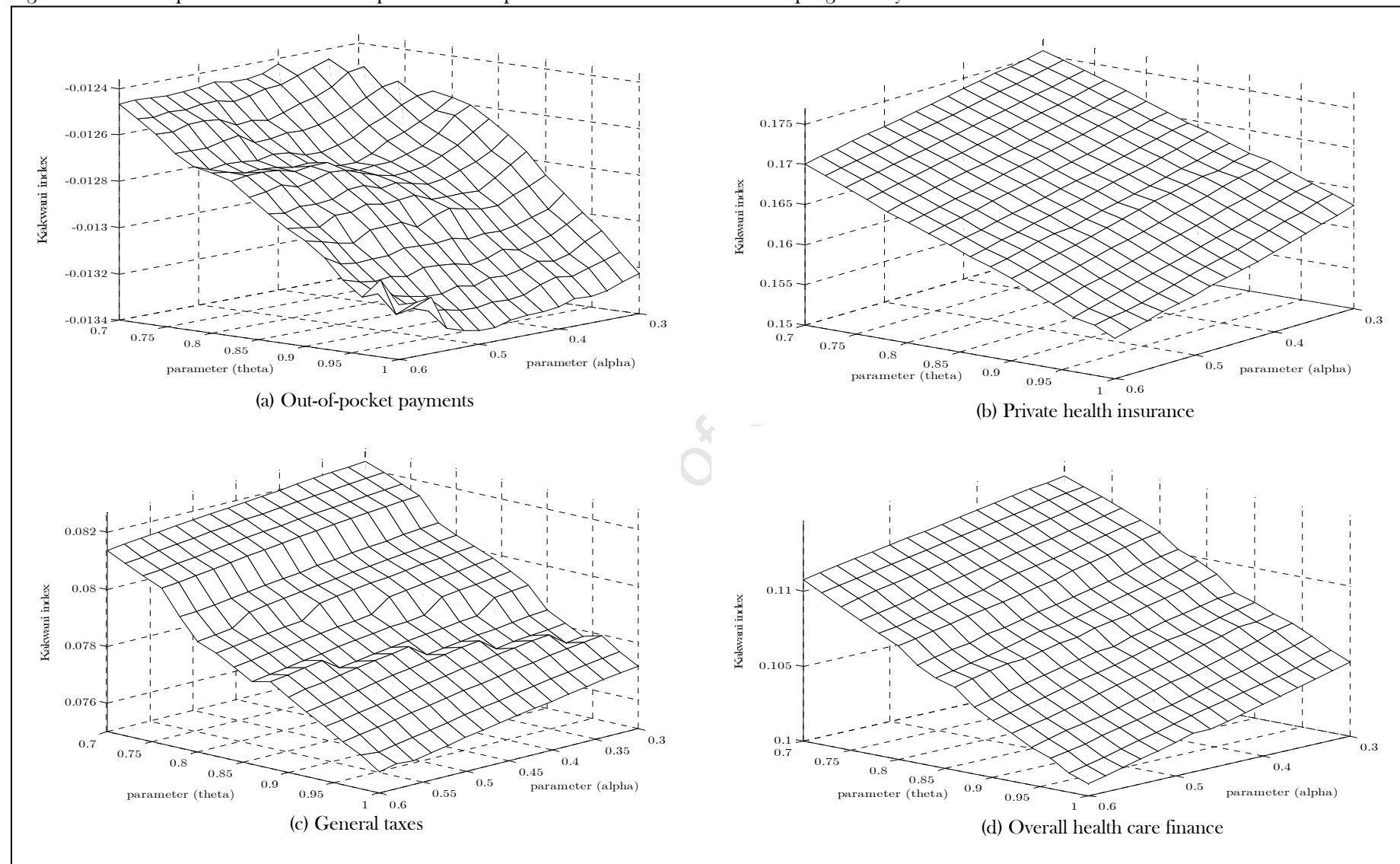
Because the answers to these questions are not particularly straightforward Figure 6.9 presents an assessment of the robustness of the Kakwani index to the choice of different equivalence scale parameters. It should be recalled that the equivalence scale used is of the form $AE = (n_A + \alpha n_K)^\theta$ where the parameters alpha (α) and theta (θ) represent the cost of children and economies of scale respectively. With the charts in Figure 6.9 we can simultaneously assess the impact of changing these two parameters on progressivity.

The charts in Figure 6.9 generally indicate that the value of the Kakwani index decreases mildly with both θ and α . This is consistent for out-of-pocket payments (Figure 6.9(a)), private health insurance (Figure 6.9(b)), general taxes (Figure 6.9(c)) and overall health care financing (Figure 6.9(d)). Though the changes in Kakwani indices are generally small, the index is more responsive to changes in θ than to changes in α . This implies that economies of scale have greater influence on the results of progressivity than the cost of children. For overall health care financing shown in Figure 6.9(d) for instance, the Kakwani index increased marginally from 0.10 (when $\theta = 1$) to 0.11 (when $\theta = 0.7$) for the value of $\alpha = 0.6$. Reducing household

economies of scale (i.e. $\theta \rightarrow 1$) leads to a reduction in the progressivity of health care financing. In the case of a regressive financing mechanism such as out-of-pocket payments, the extent of regressivity is increased. Therefore when households live more cheaply, payments for health care are likely to be more progressive than they are when there are no economies of scale. Increasing the cost of children reduces (increases) progressivity (regressivity). The trend is similar across all the charts in Figure 6.9 for any choice of θ . For instance as shown in Figure 6.9(d) when the value of $\theta = 1$ and $\alpha = 0.3$, Kakwani index of overall health care financing was 0.105 but this only decreased marginally to 0.10 when $\alpha = 0.7$.

University of Cape Town

Figure 6.9: The impact of the choice of equivalent scale parameters on the assessment of progressivity

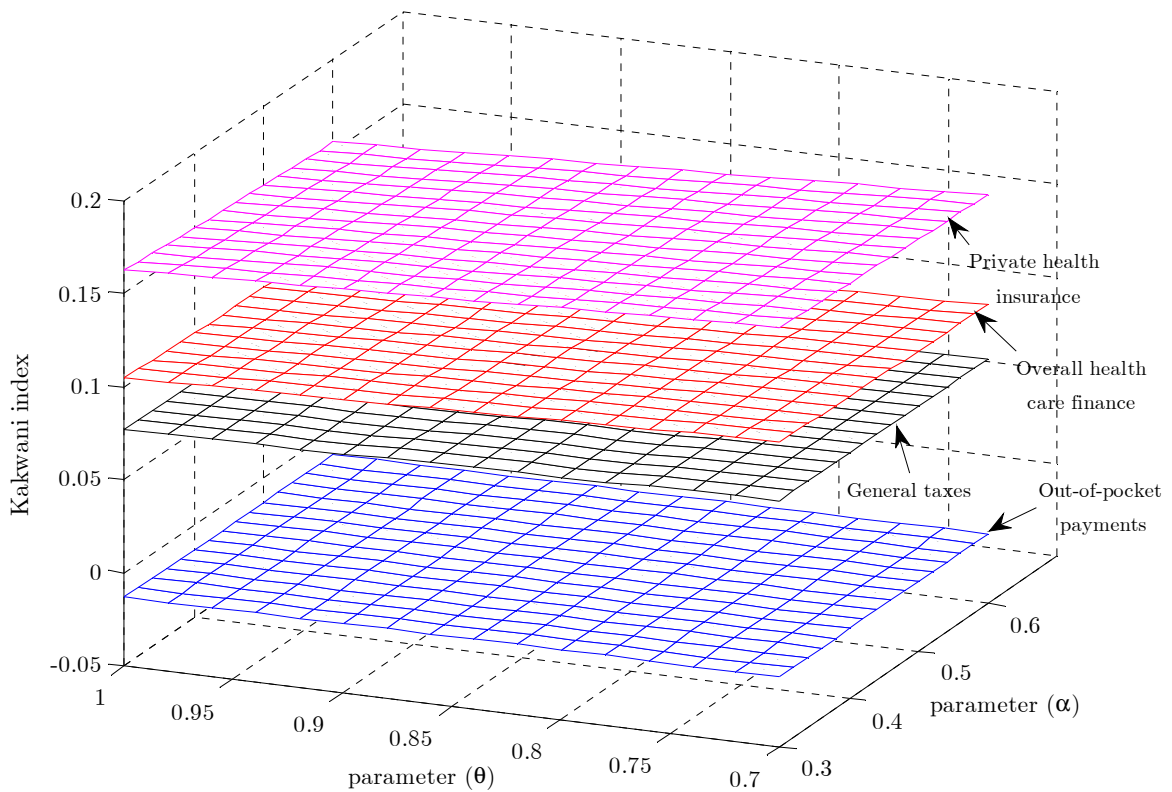


The results of these sensitivity tests show therefore that though progressivity is slightly responsive to the equivalence scales parameters, they are less sensitive to how much less or more we weight the consumption of children relative to adults. The values of economies of scale parameters in this case have greater impact on the assessment of progressivity in such a way that as families become able to achieve economies of scale, health care financing is less likely to place a greater burden on the poor compared to the rich.

Generally studies have shown that these measures of inequality, as reported in this study, are not very sensitive to the choice of equivalence scale parameters (Coulter *et al.*, 1992, Cutler and Katz, 1992, Jenkins and Cowell, 1994, Burkhauser *et al.*, 1996). Cutler and Katz (1992) for instance have shown this in the context of inequality in the United States and Burkhauser *et al.* (1996) for Germany and the United States. Buhmann *et al.* (1988) using cross country comparison however noted that there are likely changes in ranking of countries using different equivalence scales and different measures of inequality. However Karoly and Burtless (1995) conclude in the context of United States that inequality trends are insensitive to the choice of economies of scale.

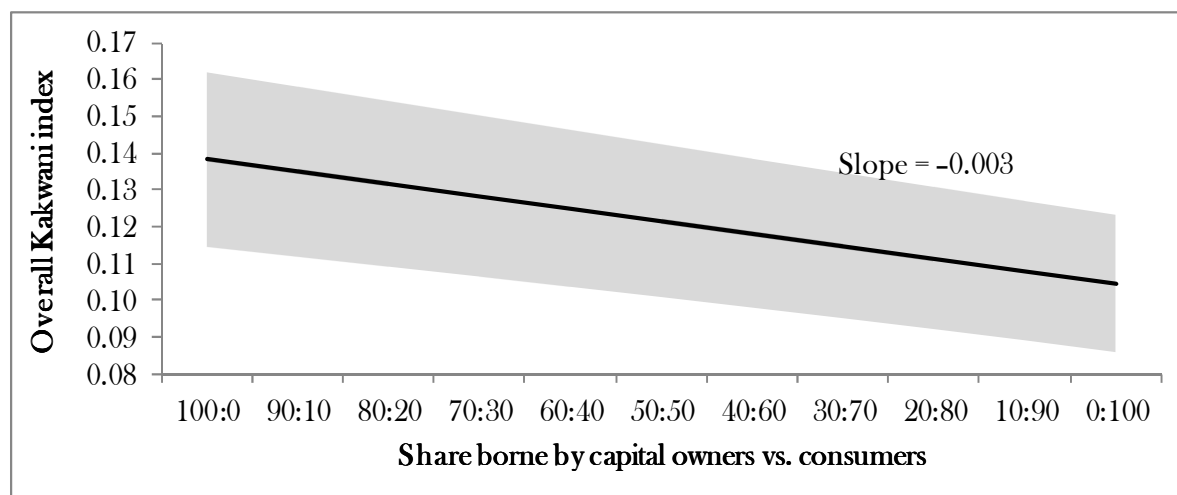
In order to investigate the likely impact of changes in the scale parameters in ranking progressivity of general taxes, health insurance, out-of-pocket payment and overall health care financing Figure 6.10 plots the sensitivity of ranking of progressivity to the choice of equivalent scale parameters for the case where $\nu = 2$. As seen in this chart, there are no inconsistent rankings due to the choice of scale parameters. Irrespective of the choice of parameter values, out-of-pocket payments are consistently regressive while the Kakwani indices for private health insurance are still the largest. Aaberge and Melby (1998) however argue that the modest responsiveness of results to choice of equivalence scale parameters may be because there are little or no differences made between consumption of the poor relative to that of the rich in assessing household needs. Even after accounting for this they found that it is only the magnitudes of inequality that consistently change but not the overall trend. The result of their assessment is therefore similar to that presented in Figure 6.8 in combination with that presented in Figure 6.10.

Figure 6.10: The impact of the choice of equivalence scale parameters in ordering progressivity ($\nu = 2$)



The analysis of the sensitivity of the progressivity of health care financing to changes in the assumptions about where the economic incidence of corporate income tax lies (i.e. the distribution between capital owners or shareholders and households as consumers) is presented in Figure 6.11. The result shows that overall progressivity is robust to assumptions made about the shifting and incidence of corporate income tax. Over the entire range of assumptions, the Kakwani index is still positive and statistically different from zero (indicating progressivity in the overall system). It however decreased mildly (slope = -0.003) from about 0.14 (when the entire burden is borne by capital owners) to 0.10 (when the burden rests solely on consumers).

Figure 6.11: The sensitivity of the Kakwani index of total health care financing to assumptions about corporate income tax



Note: 95% confidence interval is shown around the estimate

6.6 DISCUSSION

The results presented in this paper indicate that there is a fair degree of progressivity in health care financing in South Africa. The Kakwani index of overall health care financing was estimated at 0.12. Among the constituents of total financing, general taxes are progressive, private health insurance is progressive and out-of-pocket payments are proportional. Though general taxes are progressive, all indirect taxes (fuel levy, value-added tax and excise taxes) were regressive. Direct taxes (personal and corporate income taxes) were however progressive. The Kakwani index for out-of-pocket payments initially indicates a regressive relationship but this was rejected by the test of dominance and the conclusion was that out-of-pocket expenditures are proportional. This means that they place similar levels of burden on all households. The case of private health insurance, called medical schemes in South Africa, is very different. This financing mechanism was found to be progressive implying that the rich bear a disproportionately larger share of the burden. The distribution of enrolled members follows income gradients (McIntyre and Doherty, 2004, Borghi *et al.*, 2009, Ataguba and Goudge, 2012) because the poor cannot afford to purchase cover. Among the insured, however, it was found to be regressive.

As also reported in this study, international evidence shows that out-of-pocket payments are generally regressive or proportional (cf. Lairson *et al.*, 1995, Wagstaff *et al.*, 1999b, Cissé *et al.*,

2007, Ugá and Santos, 2007, O'Donnell *et al.*, 2008a, Smith, 2010). These results are consistent across high income countries (cf. Wagstaff *et al.*, 1999b, Klavus, 2001) and middle-to low- income countries (cf. Cissé *et al.*, 2007, O'Donnell *et al.*, 2008a). Exceptions include a few countries such as Bangladesh, Indonesia, Philippines and Zambia (O'Donnell *et al.*, 2008a, Mwenge, 2010) where out-of-pocket payments are clearly progressive. In some countries where a positive value of the Kakwani index is reported, the value is usually smaller than 0.07 and this shows an almost proportional relationship. The value of the Kakwani index for out-of-pocket expenditures obtained in this study (-0.002) compares with that reported elsewhere for Colombia (Castano *et al.*, 2002). The studies that have recorded positive Kakwani indices have noted that it reflects constrained financial access of the poor (Mwenge, 2010). In many developing countries out-of-pocket payments are a substantial component of overall health care expenditures (Lu *et al.*, 2009, Ataguba, 2012b) and this form of financing health services is described as inefficient and inequitable (McIntyre and Doherty, 2004, Knaul *et al.*, 2006). Inequity here derives from the non-discriminatory nature of the payments. Though user fees are differentiated by income levels in South Africa, in most countries, it is usually the case that both the rich and the poor are frequently faced with the same payment schedule independent of income level. Payments are also highly fragmented (McIntyre *et al.*, 2008) in such a way that there are no pooled risks. These create inefficiencies in the system. In the case of South Africa, as found in this study, out-of-pocket payments comprise about 14% of total health care expenditures and the distribution of the burden, as reflected in the Kakwani index, confirms the inequitable nature of such payments. This could still represent an underestimation of the burden households' face. This is because some households that cannot afford payments may have forgone treatment thereby imposing higher burden in the long-run as poorer households delay treatment until the illness becomes severe, requiring even more expensive treatment (McIntyre *et al.*, 2006).

The distribution of indirect taxes as obtained in this study conforms to those reported internationally (Wagstaff *et al.*, 1999b, Klavus, 2001, Smith, 2010, Yu *et al.*, 2011). Just like out-of-pocket payments, with the exception of some countries, indirect taxes are mostly regressive health care financing mechanisms. Only in some countries such as China, Bangladesh, Hong Kong, Indonesia, Thailand, Nepal and Philippines (see O'Donnell *et al.*, 2008a) have indirect taxes been reported as progressive. Generally in most countries where indirect taxes are progressive, out-of-pocket payments tend to also be progressive. Most of these studies that investigate indirect taxes do not show the progressivity of the constituents of

total indirect taxes. In this regard it is difficult to compare the regressive results in the components of indirect taxes obtained in this study with external studies. However the global trend is that developing countries with large informal sectors tend to record progressive indirect taxes. Because tax structures and policies differ across developed and developing countries (Gordon and Li, 2009), the extent to which tax laws are enforced differ. One of the few studies on health care financing that differentiated the components of indirect taxes was conducted in Ghana (Akazili, 2010, Akazili *et al.*, 2011). In that study indirect taxes were progressive because import duty and value-added tax are progressive and these offset the regressivity of the fuel levy. Value-added tax was progressive in Ghana because of selective exclusion of certain commodities consumed mainly by the poor from taxation (Akazili *et al.*, 2011). In the case of South Africa, as found in this study, all the components of excise taxes are regressive. The most regressive being excise taxes on alcohol and tobacco products. The regressivity of alcohol taxes in South Africa has been reported elsewhere (Ataguba, 2012a) and this is driven by the distribution of alcohol drinkers. Also value-added tax, excise taxes and fuel levy have been reported as regressive in South Africa (Simkins *et al.*, 2000, Go *et al.*, 2005, Mabugu *et al.*, 2009).

In some countries such as Ghana and Egypt some indirect taxes are earmarked specifically for the health sector and they each account for about 3% of total health care expenditures in these countries. For example Ghana earmarks 2.5% of value-added tax rate as a national health insurance levy to finance their national health insurance (Akazili *et al.*, 2011) especially to ensure that the contributors' base is increased. This financing source was found to be mildly progressive (albeit not statistically significant) in Ghana because it is essentially a reflection of the distribution of value-added tax. Egypt operates an earmarked health tax on the sale of cigarette and related products (O'Donnell *et al.*, 2008b) and the distribution was found to be proportional. These imply that countries could rely more on indirect taxes for funding health care when there is some progressivity in payments.

Direct taxes were reported as progressive health care financing sources in South Africa. This result is consistent with previous literature both in developing and developed countries. The Kakwani index of 0.21 obtained in this study compares to those reported in other studies around the world. For example, the Republic of Korea in 2000 recorded an index value of 0.27 and Egypt a value of 0.25 in 1997 (O'Donnell *et al.*, 2008a, O'Donnell *et al.*, 2008b). Most of the studies do not distinguish between personal income tax and corporate income tax

in arriving at the incidence of direct taxes (cf. Wagstaff *et al.*, 1999b, Ugá and Santos, 2007, O'Donnell *et al.*, 2008a, Yu *et al.*, 2011). Studies in Ghana and Tanzania however distinguished between corporate income tax and personal income tax and both were found to be progressive (Mtei and Borghi, 2010, Akazili *et al.*, 2011). In general, all tax incidence studies conducted in South Africa have found that direct taxes are progressive (Simkins *et al.*, 2000, Amm, 2001). Amm's (2001) study investigates corporate income tax in South Africa by focusing on shareholders, both direct and indirect (i.e. via pension fund investments). The study showed that while corporate income tax is progressive the inclusion of indirect shareholders rather than only considering direct shareholders (through dividends) reduces the extent of progressivity. In all countries where indirect taxes as a source of health care financing were reported as progressive, the value of the Kakwani index for direct taxes is greater than that for indirect taxes (see O'Donnell *et al.*, 2008a for the case of Bangladesh, Hong Kong, Indonesia, Korea Republic, Nepal, Philippines and Thailand). This signifies that even when indirect taxes are progressive, they are less progressive than direct taxes. The main reason for progressive direct taxes is the progressively structured personal income tax rates. In all the countries though there are differences in the way personal income tax rates are set; they generally discriminate favourably between the poor and the rich. In South Africa the marginal tax rate as is an increasing function of income (see Table 5.2).

Progressive direct taxes and regressive indirect taxes generally result in progressive general taxes because of the dominance of progressive direct taxes over indirect taxes. Results elsewhere in Australia, Brazil, Denmark, Finland, Germany, Ghana, Ireland, Italy, Netherlands, Portugal, Spain, Sweden, Switzerland, Tanzania, United Kingdom and USA have consistently shown that general taxes are progressive sources of health care financing (see Lairson *et al.*, 1995, Wagstaff *et al.*, 1999b, Mtei and Borghi, 2010, Smith, 2010, Akazili *et al.*, 2011). The results in this paper conform to those of previous studies. Financing health care through general taxes will therefore place lesser burden on the poor relative to the rich than through out-of-pocket payments. Countries such as the United Kingdom and more recently Thailand that finances health care largely through progressive general taxes place less burden on the poor. However countries that rely heavily on other regressive financing sources such as out-of-pocket payments and in some cases social health insurance (for example Germany and Netherlands) tend to place a greater burden of health care financing on the poor.

Private health insurance may differ across countries. Their presence in Africa and other developing countries is limited. In some countries such as Tanzania, community insurance schemes exist (Mtei and Borghi, 2010). The type of private health insurance that exists in a country and the structure of premiums will determine the distribution of the burden. In Tanzania the Community Health Fund was found to be regressive because it covers those in the informal sector and premiums are not related to income levels (Mtei and Borghi, 2010). In South Africa private health insurance (or medical scheme) contributions were found to be overall progressive. International studies have shown that the incidence of private health insurance⁶⁰ varies from progressive to regressive. Where private health insurance is taken out as a top-up cover to supplement publicly available programmes or cover such as in Portugal and the United Kingdom, it turns out as a progressive health care financing mechanism (Wagstaff *et al.*, 1999b). This is because it is only the rich that can afford to purchase such supplementary insurance to cover for some luxuries. In fact the Kakwani indices obtained for private health insurance in Portugal and the United Kingdom exceed that from general taxes. Elsewhere in countries such as the United States and previously Switzerland where private insurance is taken out as the sole cover, and enrolment cuts across the entire population, a regressive pattern was observed (Wagstaff *et al.*, 1999b). In the case of Germany where it is taken out as the sole cover but only the rich enrol, it turned out to be a progressive financing mechanism. Another country such as Palestine has reported a progressive private health insurance (Abu-Zaineh *et al.*, 2008). Again it is progressive because it is only selected rich population that purchase cover.

Overall health care financing is progressive in South Africa. The progressive financing system is driven largely by progressive general taxes and private health insurance. This result is generally encouraging given South Africa's large income inequality. Internationally it is accepted that health care financing should follow the principle of ability-to-pay (see Chapter 3 for details). Even though there is no universally accepted extent of progressivity, financing healthcare according to ability to pay means that health care financing should be progressive (Wagstaff and van Doorslaer, 2000). Results from international literature show that health care financing systems could be progressive or regressive. These results further show that the relationship does not depend on the level of development or income of the country. Some richer countries such as Japan, Germany, Denmark, Netherlands, Sweden, and Switzerland

⁶⁰ Even when community based insurance schemes are excluded.

have regressive health care financing systems while some developing countries such as Ghana, Egypt, Bangladesh, Malaysia, and Tanzania have progressive health care financing systems. Countries that record progressive financing systems rely more on progressive mechanisms such as direct taxes and to some extent social health insurance and not necessarily when general taxes are progressive. However the Republic of Korea that relies heavily on a regressive social health insurance has a regressive overall health care financing (O'Donnell *et al.*, 2008a) even in the face of progressive indirect taxes, direct taxes and out-of-pocket payments. In the case of South Africa total health care financing is progressive because general taxes and private health insurance are progressive and these together accounts for over 85% of total health care financing in the country. Such progressive results imply that health care financing places lesser burden on the poor relative to the rich. In fact as shown in Figure 6.2, the poorest 20% of the population spend about 4% of their income on health care compared to 14% by the top 20%.

Generally it is the case that fairness or equitable financing is upheld as a goal for health systems (World Health Organization, 2010). While it is challenging to operationalize this in different contexts, South Africa recognises that equity in health care financing and use of health services are important policy goals (Department of Health, 1997, 2011) that should be pursued. However, the country is still faced with several equity challenges (McIntyre and Gilson, 2002, McIntyre *et al.*, 2002, Ataguba *et al.*, 2011, Ataguba and McIntyre, 2012). The major equity challenges include the public-private mix, maldistribution of resources across geographic regions, and disparities in the allocation of public resources between socio-economic groups (see McIntyre and Doherty, 2004, Coovadia *et al.*, 2009). These challenges are heightened by high levels of income inequality, unemployment and poverty in the country (du Toit, 2002, Development Indicators, 2010).

It is important to understand in the context of South Africa, given the challenges facing the health system, the usefulness of the results presented here. This is motivated by the fact that the choices about how health care is financed have an impact on how fairly the burden of payment is distributed (World Health Organization, 2000). When health systems are financed through regressive mechanisms, they place greater burden on the poor relative to the rich (Wagstaff *et al.*, 1999b, Wagstaff and van Doorslaer, 2000). Progressive health care financing is generally preferred. These issues in the context of South Africa are dealt with, in detail, in Chapter 8. They form the basis of discussions in that chapter.

6.7 CONCLUSION

This chapter has rigorously investigated the distribution of the burden of health care financing in South Africa using standard incidence methodologies. The results presented show that while health care financing in South Africa is progressive, not all the constituent parts are progressive. Out-of-pocket payments that comprise about 14% of total health care financing are mildly regressive (with a negative Kakwani index). Statistical tests indicate that they are proportional. All indirect taxes were regressive and all direct taxes were progressive. Private health insurance was also found to be progressive but with a skewed distribution of enrolled members. Sensitivity analyses performed on the results show that the progressivity results presented are robust to the choice of equivalence scale parameters, shifting assumptions about corporate income tax and the society's or policy maker's aversion to inequality of the poor. The results are discussed and compared with those from other different country studies.

Chapter 7

DECOMPOSITION ANALYSIS OF THE INCOME REDISTRIBUTIVE EFFECT OF HEALTH CARE FINANCING IN SOUTH AFRICA

7.1 INTRODUCTION

This chapter presents an in-depth assessment of the income redistributive effect of health care financing in South Africa. In Chapter 6, only the incidence of health care payments on households relative to their living standards was explored. That analysis assumes away the presence of horizontal inequality and re-ranking that could be associated with health care financing. For instance a health care financing system may be progressive but still induce increases in inequality through differential treatments. This is likely to be the case when there is substantial re-ranking or horizontal inequities. The aim of this chapter therefore is to unpack the components of redistribution in order to identify the areas of health care financing in South Africa that drives inequity in the system. Methodologically this chapter also compares the results from competing models and draws some methodological conclusions in the case of South Africa, a country with high income inequality. Because the analyses are based on several assumptions about the values of certain parameters, robustness and sensitivity analyses were performed. The discussions in this chapter are based on the rationale of pursuing social justice and equity (Boadway and Keen, 2000) introduced in Chapter 3. They also relate to the ATP principle.

7.2 OVERVIEW OF THE METHODOLOGY

This section presents an overview of the methodology for assessing the income redistributive effect of health care financing. The extensive evolution of these methods is contained in Chapter 4. Here, in order avoid replication, only a briefly statement of the methods are provided. Where possible some more details are provided than were previously outlined.

Such details include how the method was operationalized in this thesis. Based on the review contained in Chapter 4, three competing methods were selected. The first is the traditional Aronson, Johnson and Lambert – referred to here as *AJL* methodology or model (Aronson *et al.*, 1994). The second is an ‘improvement’ over the *AJL* methodology that attempts to assess *full vertical effect* and also decompose reranking into between-group reranking (i.e. the same as the *AJL* reranking), within-group reranking and entire group reranking. This was developed by Urban and Lambert (Urban and Lambert, 2008) and is referred to here as the *UL* methodology or model. The third methodology considered is that proposed by Duclos, Jalbert and Araar (Duclos *et al.*, 2003). This methodology decomposes inequality measures based on the Atkinson-type social welfare functions by incorporating some ethical concerns. This is referred to here as the *DJA* methodology or model. The *DJA* methodology attacks one of the challenges with assessing redistributive effects (i.e. the arbitrariness involved in the choice of near-equals as opposed to ‘real’ or ‘exact’ equals when assessing classical horizontal inequity). This is driven largely by the genuine concern that in most datasets there may be no likely occurrence of exact equals such that any arbitrary categorization of equals may limit the extent to which accurate decomposition can be performed. In their formulation therefore they suggest the use of non-parametric approaches to obviate the problems inherent in categorising near-equals. In doing so horizontal equity can be assessed in the ‘classical’ sense defined as equal treatment of equals. However one of the challenges with using the *DJA* approach is the choice of parameters of aversion to inequality and uncertainty in net incomes.

It must be borne in mind that there are two classes of methods – those based on the decomposition of the Gini index (i.e. *AJL* and *UL*) and that based on the decomposition of an Atkinson-type inequality measure (i.e. *DJA*). Because of this, the results from these methods may not be directly comparable. However for methodological reasons this chapter attempts to compare the results obtained from these methods. Such comparison is not so much based on the magnitudes of the components of the decomposition but mainly based on the expected signs of the components and their relative contribution to overall redistribution.

Let us define $\mathbf{X} = [x_1, x_2, \dots, x_{M-1}, x_M]$ as a vector of gross per adult equivalent incomes and $\mathbf{T} = [t_1, t_2, \dots, t_{M-1}, t_M]$ as the vector of per adult equivalent health care payments (e.g. private health insurance). If we define $n_i = x_i - t_i$, then $\mathbf{N} = [n_1, n_2, \dots, n_{M-1}, n_M]$ is a vector of per adult equivalent net-incomes after health care payments corresponding to the vector of per adult equivalent gross incomes.

Following Reynolds and Smolensky (1977) we can define the redistributive effect of healthcare payments (RE) under the AJL and the UL frameworks as the change in inequality associated with health care payment:

$$RE = \Delta G = G_X - G_N \quad (7.1)$$

where G_X is the Gini index associated with the vector \mathbf{X} and $G_N = G_{X-T}$ is the Gini index associated with the vector \mathbf{N} .

A pro-poor redistribution occurs when inequality in the vector \mathbf{N} is less than that in \mathbf{X} such that $RE > 0$ while a negative redistribution occurs when $RE < 0$. Aronson *et al.* (1994) further decompose the redistributive effect of payments as follows:

$$RE = V - H - R \quad (7.2)$$

where V represents vertical equity, H horizontal inequity and R reranking. Most methods used to assess redistribution have adopted this convention.

These decomposition analyses require that we compute Gini and concentration indices. Except otherwise stated, in this chapter the Gini index was computed in Stata® as:

$$G = 1 - (\hat{\xi} / \hat{\mu}_X) \quad (7.3)$$

where $\hat{\xi} = \sum_{i=1}^M ((V_i)^2 - (V_{i+1})^2) / (V_1)^2 x_i$; $V_i = \sum_{h=i}^M w_h$ and $x_1 \geq x_2 \geq \dots \geq x_{M-1} \geq x_M$. The vector $\mathbf{W} = [w_1, w_2, \dots, w_M]$ represents the household sampling weights based on the sampling design. The weighted average of the vector \mathbf{X} is given as $\hat{\mu}_X$.

The concentration index for any vector $\mathbf{Z} = [z_1, z_2, \dots, z_M]$ was similarly computed in Stata® as follows:

$$C_Z = 1 - (\hat{\xi}_Z / \hat{\mu}_Z) \quad (7.4)$$

where the ranking variable is the vector \mathbf{X} , $\hat{\mu}_Z$ is the weighted average of the vector \mathbf{Z} ; $\hat{\xi}_Z = \sum_{i=1}^M ((V_i)^2 - (V_{i+1})^2) / (V_1)^2 z_i$; and $x_1 \geq x_2 \geq \dots \geq x_{M-1} \geq x_M$.

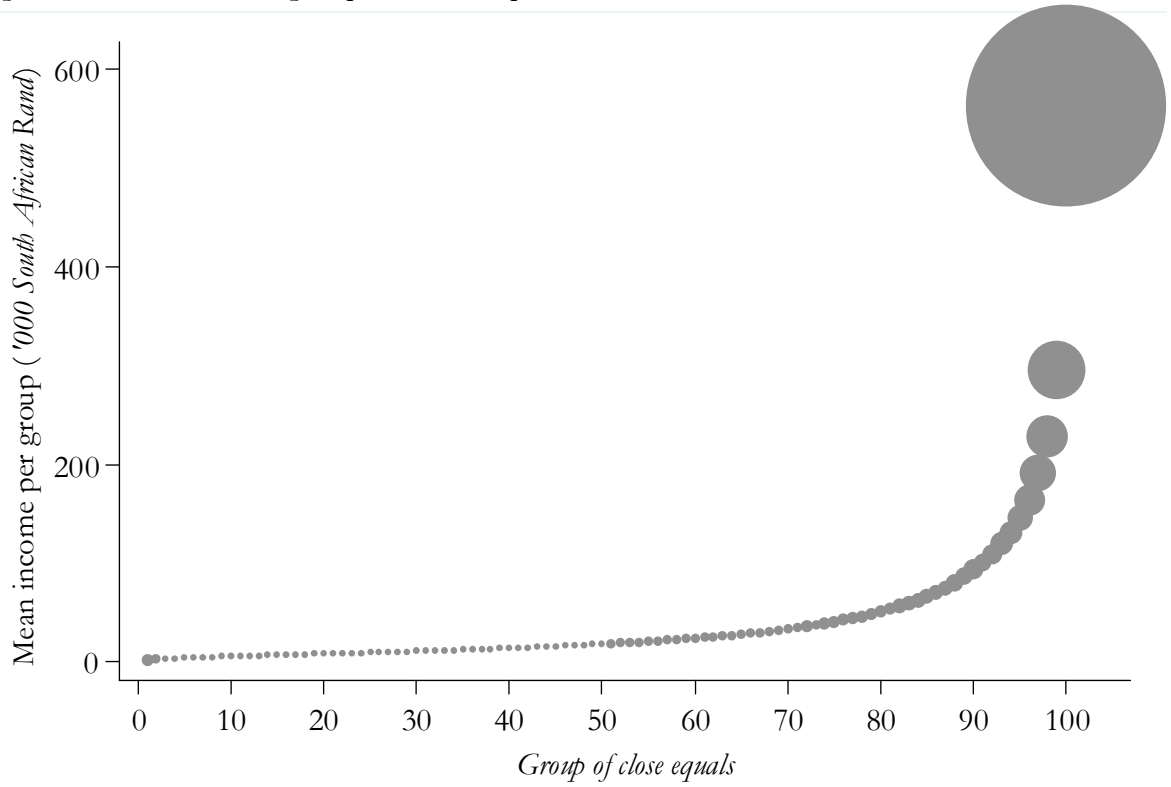
7.3 THE METHODOLOGY OF THE AJL DECOMPOSITION

The Aronson, Johnson and Lambert (Aronson *et al.*, 1994) decomposition of (7.2) can be stated as:

$$\underbrace{G_X - G_{X-T}}_{=RE} = \underbrace{(g / (1 - g))\pi_{\bar{K}}}_{=V^{AJL}} - \underbrace{\sum_j \alpha_j G_{X-T,j}}_{=H^{AJL}} - \underbrace{G_{X-T} - C_{X-T}}_{=R^{AJL}} \quad (7.5)$$

Because there are no exact equals in the dataset, based on what has been previously proposed, it requires that we first of all generate bands of equals or near equals (Aronson *et al.*, 1994). Usually most studies use predetermined fixed income bands (i.e. with equal widths) but this is hardly possible in the case of an unequal society where the distance between income points varies widely especially at the top of the distribution. Because of the large variations in the vector \mathbf{X} , the population was grouped into percentiles (i.e. 100 groups of households) based on the vector \mathbf{X} . Each percentile was treated as close-equals and used in the analysis. To understand the extent to which there are variations in the groups (especially at the top of the distribution), Figure 7.1 shows the mean adult equivalent income of each of the groups with their respective standard deviations represented by the size of the circles. Any group with a larger circle indicates that the range is wide and variations in income within the group are also high. For example the 100th group has the largest circle signifying huge variations in income at the top of the income distribution. Generally the groups below the 80th group tend to be more ‘homogenous’ than those above it.

Figure 7.1: Variations in groups of close-equals



The various components of the *AJL* decomposition in (7.5) were obtained as follows:

7.3.1 The overall redistribution

$G_X - G_{X-T}$ measures total redistributive effect of any health care financing source T . The Gini index of pre-payment income G_X was obtained directly using (7.3) while the Gini index of post-payment income G_{X-T} was also obtained using (7.3) but with \mathbf{N} as the vector of interest.

7.3.2 The *AJL* vertical component

V^{DJA} is some measure of progressivity in health care financing given that there is no reranking and there is no horizontal inequity. Here $\pi_{\bar{K}}$ is the Kakwani index obtained in the absence of horizontal inequity in health care financing. This was estimated as the difference between the between-group concentration index of health care financing (C_T^B) and the Gini index of pre-payment income (G_X). C_T^B was obtained using (6.2) where the vector of interest \mathbf{T} is replaced with group specific averages (μ_T^j) for each group j . The parameter g is the average payment or tax rate for each health care financing source (i.e. the share of payments in income).

$V^{DJA} > 0$ means that health care financing is progressive while $V^{DJA} < 0$ implies that it is regressive.

7.3.3 The AJL re-ranking component

R^{AJL} is some non-negative measure of re-ranking. Here C_{X-T} is post-payment income concentration index obtained in Stata® using (6.2) where the vector of interest is \mathbf{N} . The ranking variable is obtained by ranking households first by their prepayment income (i.e. vector \mathbf{X}) then within each group of equals, by their post-payment income (i.e. the vector \mathbf{N} for each group j) and G_{X-T} is obtained using (7.3) but the vector of interest here is \mathbf{N} .

7.3.4 The AJL horizontal component

H^{AJL} is some non-negative measure of horizontal inequity. It can be interpreted as the weighted sum of the group specific post-payment income Gini coefficients. $G_{X-T,j}$ is group j specific post-payment income Gini coefficient; α_j is the product of each group's population share and its post payment income share. In this thesis however H^{AJL} was obtained as a residual using the identity:

$$H^{AJL} = V^{AJL} - R^{AJL} - RE \quad (7.6)$$

This was done in order to reduce the computation of several Gini indices for each group. If Equation (7.6) is true then the residual is the measure of horizontal inequity.

7.4 THE METHODOLOGY OF THE UL DECOMPOSITION

Urban and Lambert's (2008) major contributions include the further decomposition of the reranking effect into three components – between-group, within-group and entire-group reranking, and the estimation of full vertical effect. The *UL* model uses the same groups of near-equals (see Figure 7.1) that were used in the *AJL* model.

The *UL* formulation can be written as:

$$\underbrace{G_X - G_{X-T}}_{=RE} = \underbrace{G_X - C'_{X-T}}_{=V^{UL}} - \underbrace{C_{X-T} - C'_{X-T}}_{=H^{UL}} - \underbrace{(R^{EG} + R^{WG} + R^{BG})}_{=R^{UL}=R^{AP}} \quad (7.7)$$

The overall redistributive effect in the *UL* decomposition in (7.7) is the same as that in the *AJL* model. The other components of the decomposition were obtained as follows:

7.4.1 The UL vertical component

V^{UL} is the vertical effect which again is some measure of progressivity. $V^{UL} > 0$ implies progressivity while $V^{UL} < 0$ implies regressivity. G_X is again obtained directly from (7.3) as the Gini index of overall income inequality. C'_{X-T} is obtained using (6.2) by replacing the variable of interest with some measure of post-payment income that deducts from each household's pre-payment income in each group j , the group specific payment share (g_j) of income.

7.4.2 The UL horizontal component

H^{UL} is a non-negative measure of horizontal inequity that captures unequal treatment of equals. Urban and Lambert (2008) calls this the 'type-2 effect'⁶¹. C_{X-T} is obtained using (6.2) with the vector of interest as \mathbf{N} and households were ranked by the vector \mathbf{X} and then by the vector \mathbf{N} among households with exact pre-payment income. C'_{X-T} was obtained as in the vertical component of the *UL*.

7.4.3 The UL re-ranking component

R^{UL} is some non-negative measure of re-ranking. Within group re-ranking R^{WG} was obtained as follows: $R^{WG} = C''_{X-T} - C_{X-T}$. While C_{X-T} is the same as that obtained in *AJL* re-ranking, C''_{X-T} was obtained using (6.2) but the variable of interest is replaced with the vector \mathbf{N} and the ranking variable is constructed by first of all ranking households within each group by their group specific vector of \mathbf{X} and then groups were ranked by group specific means of prepayment income (μ_X^j). C'''_{X-T} was obtained similarly as C''_{X-T} but with a different rank order. Here households within each group are ranked by their group vector of \mathbf{N} and groups are further ranked by their group specific mean net incomes (μ_N^j).

⁶¹ See Chapter 4 for a discussion on the type-2 effect.

Entire group re-ranking (R^{EG}) was obtained as: $R^{EG} = C'''_{X-T} - C''_{X-T}$ where C''_{X-T} and C'''_{X-T} remain as defined earlier.

Between group re-ranking (R^{BG}) is identical to R^{AJL} and it was obtained as $R^{BG} = G_{X-T} - C'''_{X-T}$ where G_{X-T} and C'''_{X-T} remain as previously defined.

Standard errors for the indices in (7.5) and (7.7) were obtained analytically using the DASP software (Araar and Duclos, 2009b). Both the sampling design features of the *IIES* 2005/06 survey discussed in Chapter 5 and the dependence of these indices were taken into account in the estimation process. However because H^{AJL} was estimated in (7.6) as the residual, analytical standard errors could not be obtained. Therefore as an alternative, as suggested in Cameron and Trivedi (2009), bootstrap methodology with 100 replications was used to obtain standard errors for H^{AJL} .

7.5 THE METHODOLOGY OF THE DJA DECOMPOSITION

As noted earlier the problem of measuring classical horizontal inequity is still inherent in the *AJL* and *UL* methods. To address this and to avoid the specification of near equal, the *DJA* approach uses social welfare functions in decomposing the redistributive effect of payments (Duclos *et al.*, 2003).

Let the vectors **X**, **T** and **N** remain as previously defined.

We can define, in a continuous sense, mean gross income as:

$$\mu_X = \int_0^1 X(p) dp \quad (7.8)$$

where $X(p)$ is the quantile (or percentile) function of per adult equivalent gross income or what Yaari (1988 p.382) loosely calls “individual p ’s income.”

If we define $N(q)$ as the quantile (or percentile) function of net per adult equivalent income (or individual q ’s net income), we can define mean net income analogously as:

$$\mu_N = \int_0^1 N(q) dq \quad (7.9)$$

We use a specification of the social welfare function by Yaari (1988) given as:

$$W_X = \int_0^1 (X(p))w(p)dp \quad (7.10)$$

where (7.10) is additive and linear in $X(p)$, the weights function $w(p)$ is defined as non-negative and depends on the position of the individual on the gross income distribution parade (i.e. it is rank-dependent). This weight function further satisfies the condition:

$$\int_0^1 w(p)dp = 1 \quad (7.11)$$

Equation (7.10) satisfies the Pigou-Dalton transfer principle and is concave in incomes only when the weights satisfy the following inequality:

$$w(p_i) \geq w(p_j) \text{ if } p_i \leq p_j \text{ for } i \neq j \quad (7.12)$$

This means that the weights attached to inequality of poorer individuals are larger than those attached to inequalities of richer individuals. This therefore means that we need to seek a specification for $w(p)$ that satisfies both (7.11) and (7.12). In this thesis the single parameter specification was used following the work of Donaldson and Weymark (1983). This is one of the most widely used specifications in the inequality literature (Duclos *et al.*, 2003). This has been used to formulate the S-Gini index or what Yitzhaki calls the extended Gini (Yitzhaki, 1983). The specification can be written as:

$$w(p;\nu) = \partial\kappa(p;\nu) / \partial p = \nu(1-p)^{(\nu-1)}, \nu \geq 1 \quad (7.13)$$

where $\kappa(p;\nu) = \nu(\nu-1)(1-p)^{(\nu-2)}$, and the single parameter ν is a measure of aversion to rank inequality (Duclos *et al.*, 2003). The larger the value of ν , the faster is the fall in $w(p;\nu)$ for higher percentiles.

To implement this with discrete data as is the case in this thesis, we write the discrete equivalence of (7.13) as (see Duclos and Araar, 2006):

$$w(p_i;\nu) = ((M-i+1)^\nu - (M-i)^\nu) / M^{\nu-1} \quad (7.14)$$

where M is the total number of households.

The specification used for the utility function is an isoelastic function (see Duclos *et al.*, 2003) given as:

$$U_{\varepsilon}(y) = \begin{cases} \frac{y^{1-\varepsilon}}{1-\varepsilon}, & \text{when } \varepsilon \neq 1 \\ \ln(y), & \text{when } \varepsilon = 1 \end{cases} \quad (7.15)$$

for $\varepsilon \geq 0$, where $U_{\varepsilon}(y)$ is a concave function with strict concavity when $\varepsilon > 0$. The function is concave because the parameter ε captures how individuals will be averse to uncertainty in their net income. ε is therefore a measure of relative aversion to risk (see Duclos *et al.*, 2003).

We redefine the social welfare function in (7.10) by replacing $X(p)$ with the utility function in (7.15) that aggregates the utilities across the entire population and taking into account the weight in (7.13). For the distribution of gross income this can be written as:

$$W_X(\varepsilon, \nu) = \int_0^1 U_{\varepsilon}(X(p))w(p, \nu) dp \quad (7.16)$$

And for the distribution of net income it can be written analogously as:

$$W_N(\varepsilon, \nu) = \int_0^1 U_{\varepsilon}(N(p))w(p, \nu) dp \quad (7.17)$$

where (7.16) is the average utility in a population weighted by the function $w(p; \nu)$ which measures relative utility deprivation in the population and (7.17) is defined analogously (see Duclos *et al.*, 2003).

It is not difficult to see that when $\varepsilon = 0$ such that $U_{\varepsilon=0}(y) = y$ (see Equation (7.15)) then Equation (7.16) becomes the S-Gini class social welfare function (Duclos *et al.*, 2003). Additionally when $\nu = 2$ this is the standard Gini social welfare function.

Following from (7.13) we know that ν is a measure of aversion to rank inequality. Yaari (1988) provides a useful way of understanding ν by defining an indicator for policy maker's degree of equality mindedness at a certain percentile p as:

$$-\frac{\kappa''(p; \nu)}{\kappa'(p; \nu)} \equiv -\frac{w'(p; \nu)}{w(p; \nu)} = (\nu - 1)(1 - p)^{-1} \quad (7.18)$$

The indicator in (7.18) is also shown in Figure 7.2. This is only dependent on the degree of aversion to rank inequality (ν) at any given rank p and not on the level of the individual's income at that rank. As shown in Figure 7.2, the larger the value of ν , the larger the degree of

equality mindedness and as shown in Figure 7.3, the faster is the fall in the weights $w(p;\nu)$ as p increases.

Figure 7.2: Policy maker's degree of equality mindedness

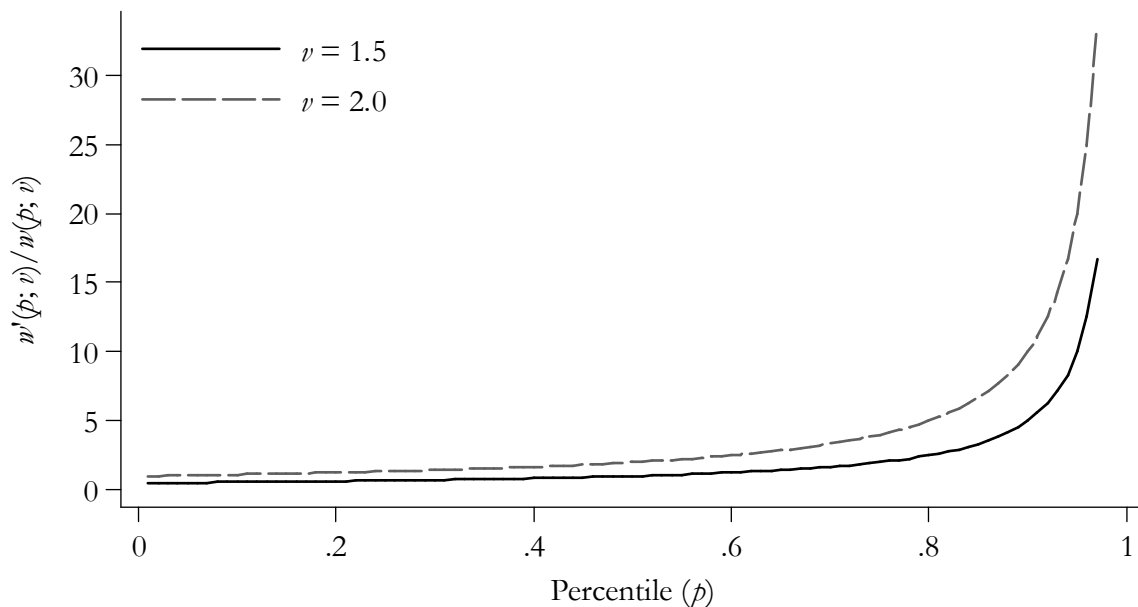
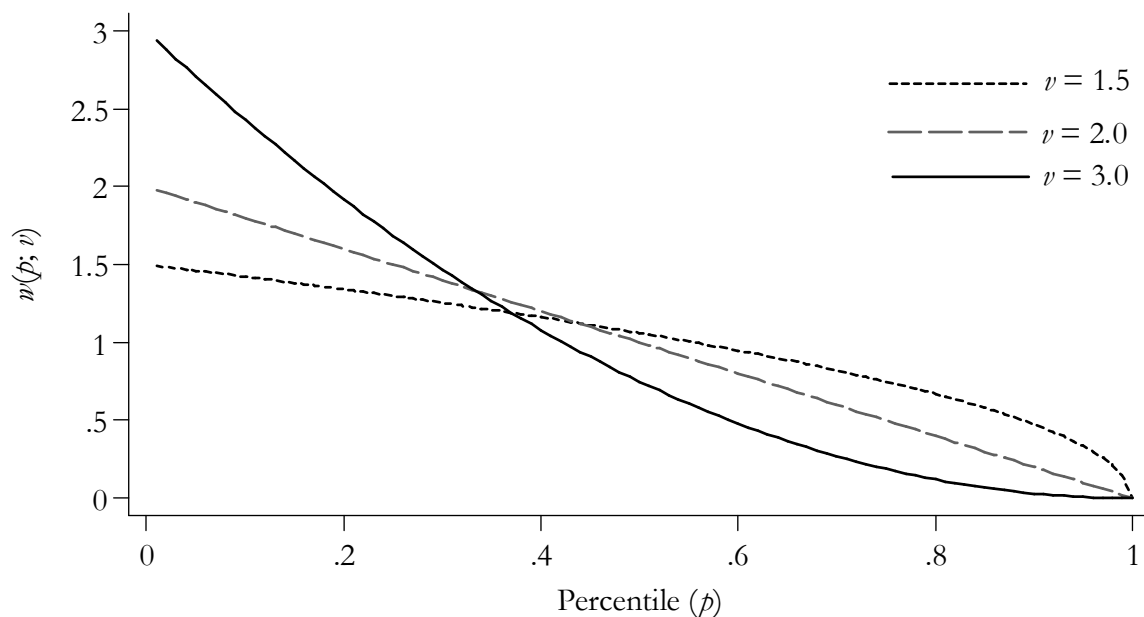


Figure 7.3 implies that for any value of ν the society or policy maker attaches more weight to inequalities or relative deprivations occurring at lower ranks than those occurring at higher ranks. Because a higher value of ν increases the society or policy maker's sensitivity to inequalities (i.e. a consequence of (7.18)), Figure 7.3 confirms that such increased sensitivity is biased in favour of the poor in aggregating relative deprivations.

Figure 7.3: The single parameter ethical weighting function ($w(p; \nu)$)



The inclusion of $w(\cdot)$ introduces relative deprivation into the assessment of inequality and moves away from the traditional utilitarian specification of the Atkinson social welfare functions (Atkinson, 1970). This is particularly important because assessment of inequality relates to how people see themselves in relation others in the society (Runciman, 1966). This has also been discussed extensively by Sen (1973, 1979, 1988).

The traditional assessment of redistribution introduced under the *AJL* and *UL* frameworks are based on the decomposition of the Gini index. The Gini index, including the extended or S-Gini is a linear index that is based on simplistic assumptions which made it popular amongst economists (Porath and Gilboa, 1994). However “the most salient drawback of linear measures is that the effect on the social welfare of a transfer of income from one individual to another depends only on the ranking of incomes but not on their absolute levels” (Porath and Gilboa, 1994 p.445). This can be seen in (7.16) when we set $\varepsilon = 0$ in (7.15) such that $U_{\varepsilon=0}(y) = y$ and $W_x(\varepsilon, \nu) = \int_0^1 (y(p))w(p, \nu) dp$ representing the weighted mean of y .

However with the non-linear representation in (7.16) and for $\varepsilon \neq 0$ the drawback in linear measures is obviated since $U_\varepsilon(y)$ is not a linear transformation of gross or net income.

If we define $N(q | p)$ as the q -quantile function for net income conditional on the p -quantile function for gross income, the expected net income of individuals at rank p in the distribution of gross income can be written as:

$$\bar{N}(p) = \int_0^1 N(p | q) dq \quad (7.19)$$

Loosely speaking $N(q | p)$ is the net income of an individual whose net income rank is q among individuals whose gross income rank is p . A non-parametric model (a Kernel regression) is used to obtain the expected net income $\bar{N}(p)$ conditional on net income quantiles. If every individual at rank p of gross income distribution were treated equally such that there is no horizontal inequity in health care payments, then social welfare function can be written as:

$$W_N^E(\varepsilon, \nu) = \int_0^1 U_\varepsilon(\bar{N}(p)) w(p, \nu) dp \quad (7.20)$$

This represents the social welfare function that is expected if there is horizontal equity such that equals were treated equally.

The expected net income utility of individuals at rank p with realised net income $N(q | p)$ is also obtained as:

$$\overline{U_\varepsilon(N(p))} = \int_0^1 U_\varepsilon(N(q | p)) dq \quad (7.21)$$

If individuals are given their expected net income utility $\overline{U_\varepsilon(N(p))}$ in place of their expected net income $\bar{N}(p)$, so also that there is some sort of local horizontal equity (Duclos *et al.*, 2003), then the resulting social welfare function can be written as:

$$W_N^P(\varepsilon, \nu) = \int_0^1 \overline{U_\varepsilon(N(p))} w(p, \nu) dp. \quad (7.22)$$

Because we use sample dataset, Equations (7.16), (7.17), (7.20) and (7.22) were estimated using the discrete formulation of $w(p; \nu)$ in (7.14) as follows:

$$\hat{W}_X(\varepsilon, \nu) = \sum_{i=1}^M U_\varepsilon(x_i) D(\nu) \text{ ordered by } x_i \quad (7.23)$$

$$\hat{W}_N(\varepsilon, \nu) = \sum_{i=1}^M U_\varepsilon(n_i) D(\nu) \text{ ordered by } n_i \quad (7.24)$$

$$\hat{W}_N^E(\varepsilon, \nu) = \sum_{i=1}^M U_\varepsilon(\hat{n}_i) D(\nu) \text{ ordered by } x_i \quad (7.25)$$

$$\hat{W}_N^P(\varepsilon, \nu) = \sum_{i=1}^M U_\varepsilon(n_i) D(\nu) \text{ ordered by } x_i \quad (7.26)$$

where $D(\nu) = ((M - i + 1) / M)^\nu - ((M - i) / M)^\nu$

By the principle of concavity and the results of (7.19), (7.20) and (7.21) it is not difficult to see that $U_\varepsilon(\bar{N}(p)) \geq \overline{U_\varepsilon(N(p))}$ and according to Duclos *et al.* (2003) it captures the local welfare cost at p of net income uncertainty. Therefore it would hold that:

$$W_N^E(\varepsilon, \nu) \geq W_N^P(\varepsilon, \nu) \quad (7.27)$$

Now if we define the equally distributed equivalent income for the distribution of gross income as $\xi_X(\varepsilon, \nu)$ such that if this is enjoyed by all individuals it would yield the same level of social welfare as the original distribution of gross income, then we can replace $\xi_X(\varepsilon, \nu)$ in (7.16) to obtain:

$$W_X(\varepsilon, \nu) = \int_0^1 U_\varepsilon(\xi_X(\varepsilon, \nu)) w(p, \nu) dp = U_\varepsilon(\xi_X(\varepsilon, \nu)) \quad (7.28)$$

This result is true because $U_\varepsilon(\xi_X(\varepsilon, \nu))$ does not depend on p and the identity in (7.11) holds.

If $W_X(\varepsilon, \nu) = U_\varepsilon(\xi_X(\varepsilon, \nu))$ and the inverse of the utility function given in (7.15) can be written as:

$$U_\varepsilon^{-1}(y) = \begin{cases} (1 - \varepsilon)y^{1-\varepsilon}, & \text{when } \varepsilon \neq 1 \\ \exp(y), & \text{when } \varepsilon = 1 \end{cases} \quad (7.29)$$

then we can write out the following: $\xi_X(\varepsilon, \nu) = U_\varepsilon^{-1}W_X(\varepsilon, \nu)$; $\xi_N(\varepsilon, \nu) = U_\varepsilon^{-1}W_N(\varepsilon, \nu)$; $\xi_N^E(\varepsilon, \nu) = U_\varepsilon^{-1}W_N^E(\varepsilon, \nu)$ and $\xi_N^P(\varepsilon, \nu) = U_\varepsilon^{-1}W_N^P(\varepsilon, \nu)$.

After obtaining all these measures we can use the general notation for inequality (see Duclos *et al.*, 2003) following Atkinson (1970) to obtain the following:

$$I_X = 1 - \frac{\xi_X(\varepsilon, \nu)}{\mu_X} \quad (7.30)$$

$$I_N = 1 - \frac{\xi_N(\varepsilon, \nu)}{\mu_N} \quad (7.31)$$

$$I_N^E = 1 - \frac{\xi_N^E(\varepsilon, \nu)}{\mu_N^E} \quad (7.32)$$

$$I_N^P = 1 - \frac{\xi_N^P(\varepsilon, \nu)}{\mu_N^P} \quad (7.33)$$

Inequality measure in (7.30) for instance can be interpreted as the percentage of total income that could be spent in removing inequality without any loss in social welfare (Duclos *et al.*, 2003).

Based on the preceding, and on the general definition of income redistributive effect, the redistributive effect of health care financing under the *DJA* formulation can be written as the change in inequality.

$$\Delta I = I_X - I_N \quad (7.34)$$

Note that the mean of net incomes contained in (7.31), (7.32) and (7.33) are identical such that: $\mu_N = \mu_N^E = \mu_N^P$ (Duclos *et al.*, 2003). Based on (7.27) we can write the following inequality:

$$I_N \geq I_N^P \geq I_N^E \quad (7.35)$$

Duclos *et al.* (2003) decomposes (7.34) taking note of the inequality in (7.35) as:

$$I_X - I_N = \underbrace{(I_X - I_N^E)}_{V^{DJA}} - \underbrace{(I_N^P - I_N^E)}_{H^{DJA} \geq 0} - \underbrace{(I_N - I_N^P)}_{R^{DJA} \geq 0} \quad (7.36)$$

As before, the extent of vertical inequity is measured by V^{DJA} and it represents the extent to which health care payments treats unequals unequally (Duclos *et al.*, 2003). Horizontal inequity is measured by H^{DJA} which represents the extent to which health care payments treats pre-payment equals unequally. In this case H^{DJA} measures classical horizontal inequity. This is because it does not require the specification of income bands that is required under the *AJL* or *UL* models. The measure of re-ranking R^{DJA} represents the extent to which individuals are overtaken by others or the extent to which they overtake others in the redistributive process. It can also be interpreted as “improper treatment of unequals” (Abu-Zaineh *et al.*, 2009). Duclos and Araar (2006 p.147) summarises this by saying: “we may also think of individuals resenting being outranked by others, but enjoying outranking others, and then assess their net feeling of resentment by the amount by which the net income of the richer (than themselves) actually exceeds what the net income of the richer class would have been had no ‘new rich’ displaced ‘old rich’ in the distribution of net incomes.”

The major challenges however with the *DJA* approach lie in the choice of parameters of aversion to inequality (ν) and aversion to risk or uncertainty in net income (ε). As discussed in Chapter 4, empirical values for these parameters have been based on a “leaky bucket” or efficiency loss experiment (Duclos, 2000) that assesses the extent of society’s tolerance to costs incurred when transferring income from a rich to a poor individual (see Duclos *et al.*, 2003). Based on this experiment the values of ε should range between 0.25 and 1.0 while that of ν should be between 1 and 4 (Duclos, 2000, Duclos *et al.*, 2003). However Duclos *et al.* (2003) note that ‘reasonable’ values for ν and ε are 1.5 and 0.4 respectively. These reasonable values have been used in recent studies on health care financing (for example see Bilger, 2008, Cavagnero and Bilger, 2010). In this thesis, except otherwise stated, the same values were used for the parameters but later perform sensitivity analyses using the range of values implied by the leaky bucket experiment.

Bootstrap methods (Efron and Tibshirani, 1986, Efron, 1987) were used to obtain standard errors for RE^{DJA} , V^{DJA} , H^{DJA} and R^{DJA} using 100 replications (see StataCorp, 2009). To avoid inconsistent estimates of bootstrap standard errors, bootstrap resamples needed to take into account the full sampling structure. This includes the stratification and clustering in the original dataset and the non-parametric nature of some estimates (Cameron and Trivedi, 2009) especially the Kernel estimations for the indices. DAD⁶² software (Duclos *et al.*, 2010) had been programmed to take all these into account and was used to obtain all estimates for the *DJA* approach.

7.6 RESULTS OF EMPIRICAL ASSESSMENT OF THE REDISTRIBUTIVE EFFECT OF HEALTH CARE FINANCING IN SOUTH AFRICA

This section presents the results of the redistributive effect of health care payments based on the *DJA* model. These results were later on compared with those obtained from the *AJL* and *UL* methodologies.

The results presented in Table 7.1 indicate that financing health care via personal income tax in South Africa reduces inequality in income. The redistributive effect (0.00430) was

⁶² This software is freely available and it is used for distributive analysis. I am grateful to the authors particularly Abdelkrim Araar and Jean-Yves Duclos for technical assistance during estimation. The software can be downloaded from: <http://132.203.59.36/DAD/index.html>.

statistically significant. The vertical effect was estimated at 0.00431. This was significantly different from zero and it indicates the extent of equity in financing health care through personal income tax. The positive value of the estimate implies that personal income tax is progressive in the sense that unequals are treated unequally such that the poor are favourably treated. This effect represents the amount by which inequality in post payment incomes could decrease if there are no differential treatments arising from reranking or horizontal inequity. A very small and negligible horizontal effect was estimated (0.000005). This was not statistically different from zero at conventional levels and it is not surprising because personal income tax in South Africa is progressively structured to treat equals equally. Reranking was estimated at 0.00001 and this was statistically different from zero. Theoretically this means that some individuals are being outranked by others in the redistributive process. However in the current context, it is important to note that the assessment of redistribution was based on adult equivalent income which is however not the basis for personal income taxation in South Africa. It is not unlikely that some households that earn the same amounts could have large household size and as such be ranked lower than their colleagues when household size and composition is taken into account.

Table 7.1: *DJA* decomposition of the income redistributive effect of health care payments, South Africa 2005/2006

Finance source	RE^{DJA}	V^{DJA}	H^{DJA}	R^{DJA}
Personal income tax	0.4299*** (0.04150)	0.4314*** (0.04142)	0.0005 (0.00047)	0.0010*** (0.00009)
Corporate income tax	0.0380 (0.02383)	0.0543** (0.02600)	0.0049*** (0.00183)	0.0114*** (0.00333)
Direct taxes	0.4717*** (0.04126)	0.4897*** (0.04156)	0.0054*** (0.00196)	0.0127*** (0.00342)
Value-added tax	-0.0576*** (0.01097)	-0.0571*** (0.01098)	0.0000 -	0.0005*** (0.00003)
Excise taxes	-0.0425*** (0.00310)	-0.0416*** (0.00316)	0.0005*** (0.00016)	0.0004*** (0.00004)
Fuel levy	-0.0204*** (0.00234)	-0.0202*** (0.00234)	0.0001 (0.00007)	0.0000 -
Indirect taxes	-0.1220*** (0.01430)	-0.1202*** (0.01443)	0.0007* (0.00042)	0.0011*** (0.00008)
General taxes	0.3546*** (0.03687)	0.3759*** (0.03729)	0.0066*** (0.00201)	0.0148*** (0.00375)
Private health insurance	0.9401*** (0.12467)	1.1091*** (0.12942)	0.0694*** (0.00729)	0.0996*** (0.00997)
Out-of-pocket payments	-0.0574** (0.02889)	-0.0297 (0.03032)	0.0124*** (0.00199)	0.0153*** (0.00320)
Overall health care payments	1.2738*** (0.13699)	1.5499*** (0.15421)	0.1076*** (0.00941)	0.1685*** (0.01992)

*, **, *** statistically significant at 10%, 5% and 1% levels respectively

The values of *DJA* parameters are: $\varepsilon = 0.4$, $\nu = 1.5$

All figures are multiplied by 100 to enhance readability.

Financing health care through corporate income tax in South Africa induces a positive but statistically insignificant income redistributive effect. As shown in Table 7.1 it reduces inequality in income by about 0.0004. This redistributive effect is smaller than that obtained from personal income tax. The vertical effect also indicates a statistically significant progressive relationship which implies that unequals are treated unequally, with poorer groups treated more favourably. The magnitude of the vertical effect (0.0005) is larger than that of horizontal inequity (0.00005) and re-ranking (0.0001). Though personal income tax is horizontally equitable, corporate income tax is horizontally inequitable. This inequity in corporate income tax is statistically significant and implies that equals are not being treated equally. This result is because corporate income tax is paid not only by owners of capital and labour but also consumers. Because both the rich and the poor consume manufactured products, and there is no explicit differentiation in prices, horizontal inequity and reranking is likely to occur. Some

of the horizontal effects may not be interpreted as inequitable because they reflect choices that individuals make but overall the lack of differentiation in those who contribute to health care financing via corporate income tax induces differential treatments that cause horizontal inequity and reranking.

Overall, financing health care through direct taxes induces significant pro-poor income redistribution in South Africa. The redistributive effect associated with direct taxes was estimated at 0.0047 which implies that inequality is reduced through direct income taxes in South Africa. This positive and significant income redistributive effect resulted from a positive vertical effect that dominates the negative effects of reranking and horizontal inequity. Because corporate income tax and personal income tax exhibit significant pro-poor vertical effects, direct taxes are also significantly progressive. The vertical effect associated with financing health care in South Africa through direct taxes was estimated at 0.0049 and this is also greater than the combined negative effects of horizontal inequity and reranking. This vertical effect is larger than the individual vertical effects of personal income tax and corporate income tax. This is because it is the weighted sum of these two. Significant differential treatments in the form of horizontal inequity and re-ranking were recorded for direct taxes largely because of the prominence of these effects with corporate income tax. Horizontal inequity in direct taxes is 0.00005 while re-ranking is 0.00013.

While financing health care through direct taxes generally leads to reductions in income inequality, indirect taxes are the exact opposite. Health care financing through value-added tax increases inequality in income. The negative redistributive effect implies that post-payment inequality increased by 0.00058 over that of pre-payment income. The major driver of the negative income redistributive effect is the pro-rich vertical effect. There is an unfavourable treatment of unequals such that the poor are bearing a greater burden of health care payments through value-added tax than the rich. Vertical inequity was significantly estimated at -0.00057. Differential treatments in the form of horizontal inequity and reranking were estimated to be very small. In fact horizontal inequity was estimated to be <0.0000001 . Reranking effect (0.000005) was however statistically different from zero. This means that when households finance health care via value-added tax, some are overtaken by others in the redistributive process.

Health care financing through excise taxes in the form of tobacco taxes and alcohol taxes redistribute income in favour of the rich. A statistically significant negative redistributive effect

(-0.00043) was estimated for excise taxes. The increase in income inequality associated with financing health care with excise tax revenue is largely associated with the significant pro-rich vertical effect. Vertical inequity was estimated at -0.00042 which means that unequal households are unfavourably treated. The poor are paying more as a portion of their income in total excise tax revenue than the rich. Differential treatments were almost equally shared between reranking and horizontal inequity. Horizontal inequity and reranking were respectively estimated at 0.000005 and 0.000004. Equal households end up being significantly treated unequally in contributing to excise tax revenue in South Africa. This is mainly because such taxes are as a result of consumption patterns of individuals and households. Only individuals that purchase items that are subject to excise taxes end up paying such taxes. The underlying differential treatment is therefore, to an extent, a reflection of the consumption decisions of individuals and households and may not necessarily be labelled as inequity. This is again largely because excise taxes are levied on items that are not necessarily categorised as necessities. These taxes are also often called sin taxes because they are levied on 'unhealthy' items (O'Donoghue and Rabin, 2006).

A fuel levy unlike alcohol taxes and tobacco taxes are not regarded as sin taxes. However just like other excise taxes, fuel levy revenue used in financing health care also significantly induces negative redistribution by increasing income inequality. The income redistributive effect was estimated at -0.0002. The extent of increases in income inequality associated with fuel levy is less than that observed for value-added tax and excise taxes. The pro-rich redistribution in fuel levy is mainly due to regressive fuel levy revenue. A statistically significant pro-rich vertical effect was estimated for fuel levy which indicates the extent to which unequal treatment of unequals increases income inequality. The reranking effect was negligible and horizontal inequity estimated at 0.000001 was not statistically different from zero.

The combination of all indirect taxes replicates the significant pro-rich redistribution associated with its components. The redistributive effect of indirect taxes was estimated at -0.0012. This is statistically different from zero and it indicates that financing health care via indirect tax revenue in South Africa increases income inequality. The increase in income inequality associated with indirect taxes (-0.00122) is less (in absolute terms) than the decrease in income inequality caused by direct taxes. As expected indirect taxes result in pro-rich vertical effects. Increase in income inequality is mainly as a result of unequal and unfavourable treatment of unequals. A statistically significant vertical inequity for indirect taxes was estimated at -0.00120.

This implies a significant regressive pattern for indirect taxes. Differential treatments associated with indirect taxes are mainly due to significant reranking effects than horizontal inequity. Horizontal inequity was marginally significant at the 10% level but the reranking effect was estimated at 0.00001 and this was significantly different from zero signifying significant switching of ranks associated with financing health care through indirect taxes.

Combining direct and indirect taxes as shown in Table 7.1, financing health care through general taxes leads to a significant reduction in income inequality. The significant redistributive effect of 0.0035 associated with general taxes is a combination of the dominant pro-poor redistribution associated with direct taxes ($RE = 0.0047$) and a dominated pro-rich redistribution associated with indirect taxes ($RE = -0.00122$). Therefore health care financing via general taxes redistributes income away from the rich towards the poor. Because $V^{DJA} > 0$ for general taxes, it means that general taxes have a pro-poor vertical effect ($V = 0.0038$). This is so because the pro-poor vertical effect for direct taxes dominates the pro-rich vertical effect observed in indirect taxes. Therefore general taxes absorb a smaller share of poor households' income compared to the share of richer households' income they absorb. The contribution of the vertical effect to redistribution is decreased by significant reranking and horizontal effects. General taxes have horizontal and reranking effects that are mainly attributed to corporate income tax and indirect taxes. Horizontal inequity was statistically significant and was estimated at 0.00007. Also reranking ($R = 0.00015$) was significantly different from zero. These mean that in the process of income redistribution through general taxes, equals were not treated equally and some households were unfavourably overtaken by others.

Pro-poor redistribution is also observed for private health insurance. Private health insurance accounts for about 45% of total health care financing and as shown in Table 7.1, contributions toward private health insurance reduce income inequality in South Africa. A significant redistributive effect ($RE = 0.0094$) was computed for private health insurance in South Africa. This significant income redistributive effect was expected because private insurance cover is purchased mainly by the rich who can afford it. The share of private insurance contributions in poor households' income is very negligible. In fact the income redistributive effect of private health insurance is far more than that resulting from general taxes. Therefore the redistribution observed here is mainly occurring among the rich who are scheme members. The positive redistribution is largely attributed to the significant positive vertical effect. The pro-poor vertical effect ($V = 0.0111$) indicates that private health insurance favourably treats unequals

unequally. Though this is the case, there is large reranking associated with private health insurance as some households, mainly those at the top of the income distribution, are outranked by their fellows. The reranking effect was estimated at 0.0010 and this was statistically different from zero. Similarly private health insurance treats equals unequally. The unequal treatment of equals gave rise to some of the re-ranking observed. A significant horizontal inequity effect was estimated at 0.0007. This horizontal inequity could be associated with the relatively similar premiums that households face. In fact some households chose less comprehensive insurance cover to mitigate the effect of the fixed premiums. However as shown in Table 7.1 these strategic decisions have not significantly reduced horizontal inequity and re-ranking of households.

Out-of-pocket payments that represent about 14% of total health care finance mildly increases income inequality. The redistributive effect of -0.00057 was computed and this is statistically significant at the 5% level. The vertical effect, though negative, was not statistically significant. However reranking and horizontal inequity were significantly associated with out-of-pocket payments. It may be argued in the context of out-of-pocket payments that the redistributive effects observed are related to the choices of individuals. However, as explained elsewhere, this is not always the case. The poor, even though the relationship is not statistically significant, pay more as a proportion of their income out-of-pocket for health care. If it is a matter of choice, we should expect that on average the rich, who can afford to pay, pay more as a proportion of their income.

Overall health care financing in South Africa significantly redistributes income to the poor. This is because as shown in Table 7.1, a reduction in income inequality is associated with financing health care. The overall redistributive effect of health care financing in South Africa was estimated at 0.01274. The pro-poor redistribution in overall health care financing is the combined effect of pro-poor income redistribution in general taxes and private health insurance, and the pro-rich redistribution in out-of-pocket payments. The increase in inequality associated with out-of-pocket payments (RE = -0.00057), though significant, was small relative to the decreases induced by general taxes (RE = 0.0035) and private health insurance (RE = 0.0094). The resultant effect is a significant pro-poor redistributive effect of total health care financing. Total health care financing was also significantly progressive. The vertical effect was estimated at 0.0155 and was significantly different from zero. This effect shows the extent to which income inequality is reduced by unequal but favourable treatment of unequals.

Even though total health care financing has a pro-poor vertical effect there were significant increases in inequality caused by differential treatment. Horizontal inequity and reranking were significant. Some households were outranked by others in the redistributive process and also equals were not treated equally. Horizontal and reranking effects were estimated at 0.0011 and 0.0017 respectively. These mean that health care financing is not entirely equitable judging from the progressivity indicator alone. The effects of differential treatment are also substantial and constitute inequity in health care financing in South Africa.

In Table 7.2 the relative contributions of the different components to overall redistribution is presented. This represents another way of presenting the results to show the major drivers of income redistribution (Wagstaff and van Doorslaer, 1997). The relative contributions of vertical equity, horizontal inequity and reranking in personal income tax to changes in inequality as presented in Table 7.2 show that vertical effect dominates. Personal income taxes would have been 0.35% more redistributive in the absence of differential treatment. Of the 0.35% effect due to differential treatment, most (0.23%) is attributed to reranking while only <0.2% is attributed to horizontal inequity. It is important to also recall that as shown in Table 7.1 horizontal and reranking effects were statistically negligible. This is expected *a priori* because of the structure of income tax rates. However this result reflects the use of per adult equivalent incomes on a household basis while income tax in South Africa is not based on households but individual income recipients. For corporate income tax, even though the redistributive effect was shown to be statistically insignificant, the vertical effect dominates. As shown in Table 7.2 corporate income tax would have been 43% more redistributive in the absence of differential treatment. About 13% of the differential treatment is due to horizontal inequity or unequal treatment of equals while a larger proportion (30%) is attributed to reranking caused by improper treatment of unequals. Direct taxes as shown in Table 7.1 were found to reduce income inequality. This reduction is mainly attributed to a dominant vertical equity effect. As shown in Table 7.2 in the absence of differential treatment, direct taxes would have been 3.8% more redistributive. About 2.7% of this differential effect is caused by reranking of households while about 1% is caused by unequal treatment of equals.

Table 7.2: *DJA* decomposition of the income redistributive effect of health care payments, South Africa 2005/2006

Finance source	(<i>V/RE</i>)%	(<i>H/RE</i>)%	(<i>R/RE</i>)%
Personal income tax	100.35%	0.12%	0.23%
Corporate income tax	142.89%	12.89%	30.00%
Direct taxes	103.82%	1.14%	2.69%
Value-added tax	99.13%	0.00%	-0.87%
Excise taxes	97.88%	-1.18%	-0.94%
Fuel levy	99.02%	-0.49%	-0.48%
Indirect taxes	98.52%	-0.57%	-0.90%
General taxes	106.01%	1.86%	4.17%
Private health insurance	117.98%	7.38%	10.59%
Out-of-pocket payments	51.74%	-21.60%	-26.66%
Overall health care payments	121.68%	8.45%	13.23%

Unlike direct taxes shown in Table 7.1, indirect taxes generally induce significant negative income redistribution. For value-added tax, the pro-rich redistribution (-0.00058) would have been 0.87% less in the absence of differential treatment. Differential treatment in the form of reranking therefore increased the extent of pro-rich redistribution. The pro-rich redistribution in financing health care through excise taxes would have been 2.12% less in the absence of differential treatment. Of this difference, the majority (1.18%) is associated with horizontal inequity and only 0.94% is due to reranking. Similarly for the fuel levy, because reranking and horizontal effects were negligible, in the absence of differential treatment the pro-rich redistribution would have been about 1% less. About 0.5% of this is due to horizontal inequity while a negligible amount is due to reranking. Income redistribution associated with indirect taxes would have been 1.5% less pro-rich in the absence of differential treatment. About 0.9% of this is attributed to reranking while 0.6% is due to unequal treatment of equals.

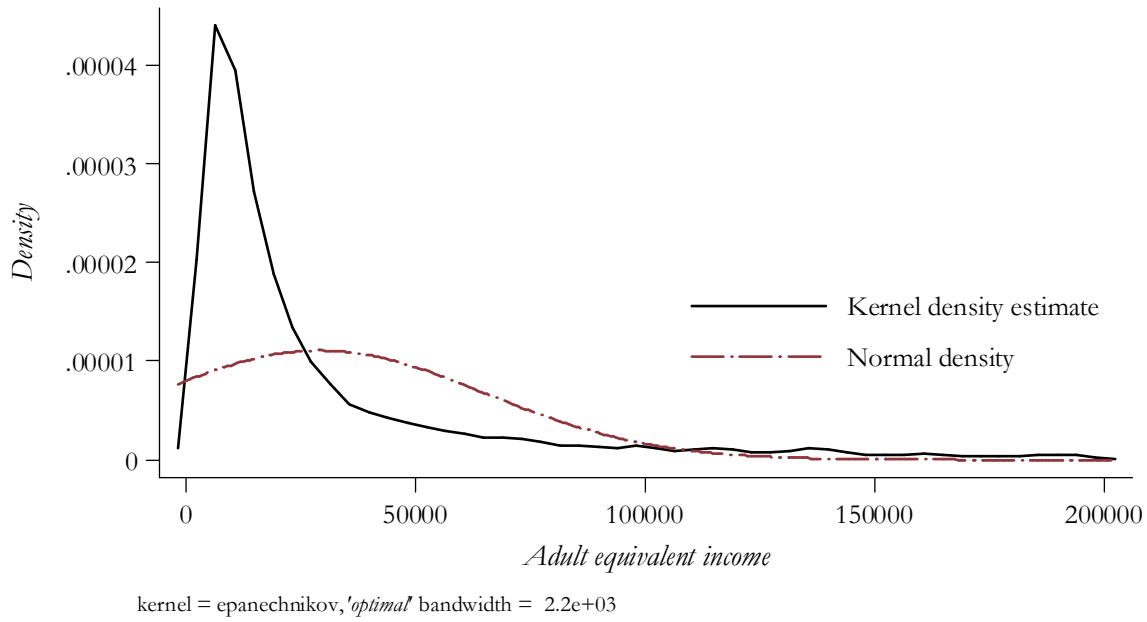
General taxes were shown to induce pro-poor redistribution. As shown in Table 7.2 vertical equity dominates. General taxes would have reduced inequality by a further 6% in the absence of differential treatments. The majority (4%) of this effect comes from reranking of households while about 1.9% comes from horizontal inequity. Similarly for private health insurance about

18% further reduction in inequality would have been observed in the absence of differential treatments of households in the redistributive process. Though the vertical component also dominates, 10.6% of the differential treatment is accounted for by reranking of households or households being outranked by others and 7.4% is due to horizontal inequity in private health insurance. Out-of-pocket expenditures, as shown in Table 7.1 however increase income inequality. As shown in Table 7.2 the increase in income inequality would have been 48% less in the absence of differential treatment of households. Reranking alone accounts for about 27% of this effect while horizontal inequity accounts for 22%. Therefore if out-of-pocket payments were 'favourable' the negative redistribution would have reduced substantially. This is similarly the case for private health insurance. All these effects put together as shown in Table 7.2 indicate that overall health care financing that induces pro-poor redistribution could have been 22% more redistributive in the absence of differential treatment of households. Over 13% of the differential treatment effect in overall health care financing is attributed to reranking of households. The dominance of reranking over horizontal inequity is expected because for general taxes, private health insurance and out-of-pocket payments the reranking effect dominates the horizontal effect. Generally in the absence of horizontal inequity overall healthcare financing would have been 8.45% more redistributive while it could have been 13.23% more redistributive in the absence of reranking.

7.7 A COMPARATIVE ANALYSIS OF INCOME REDISTRIBUTIVE EFFECTS USING THE AJL, UL AND DJA MODELS

This section compares the decomposition results of the income redistributive effect of health care financing in South Africa using the three models - *AJL*, *UL* and *DJA*. It is important to mention that in the context of South Africa where inequality in income is wide, the specification of close income groups proved very difficult (see Figure 7.1). This difficulty arises because of the skewed distribution of incomes. As shown in Figure 7.4 the distribution of income is highly skewed to the right. Among those at the top of the distribution, income is sparsely distributed. This means that the use of fixed income intervals will not be appropriate. The process of creating close-equal groups had been discussed earlier. Basically 100 groups of close equals (percentiles) were formed and used in the analysis of income redistribution under the *AJL* and *UL*.

Figure 7.4: Density plot of income distribution in South Africa, 2006



As noted earlier the results from *AJL* and *UL* are easily comparable because they are based on the decomposition of the Gini index. These results are however not directly comparable to those from *DJA* because *DJA* is based on the Atkinson-type social welfare function. However as also noted, the specification used for the social welfare functions based on Yaari (1988) allows for some generalization to the Gini social welfare function provided that the parameter of aversion to risk or uncertainty in net incomes is zero ($\varepsilon = 0$). If we set $\varepsilon = 0$, $H^{DJA} = 0$ and the only contributor to reduction in income redistributive effect becomes R^{DJA} . This in the *AJL* and *UL* models requires that the income bands shrink to zero. Note that with *AJL* the choice of income bands does not significantly affect the combined effects of reranking and horizontal inequity but only their relative contributions. Again under the *AJL* model as income bands (or intervals) are increased, reranking shrinks and when income bands are reduced horizontal inequity reduces (Aronson *et al.*, 1994). However for the *UL* model the overall reranking effect does not depend on the choice of income intervals. This can be inferred directly as a consequence of (7.7) where $R^{UL} = G_X - C_{X-T}$. Because of these issues with competing models this section does not attempt to compare the magnitudes of effects but rather the expected signs, statistical significance and relative contributions of each component to overall redistribution.

Table 7.3: Income redistributive effect of direct taxes for health care financing, South Africa 2005/2006

	Personal income tax			Corporate income tax			Direct taxes		
	<i>AJL</i>	<i>UL</i>	<i>DJA</i>	<i>AJL</i>	<i>UL</i>	<i>DJA</i>	<i>AJL</i>	<i>UL</i>	<i>DJA</i>
<i>RE</i>	0.3847*** (0.0365)	0.3847*** (0.0365)	0.4299*** (0.04150)	0.0345* (0.0190)	0.0345* (0.0190)	0.0380 (0.02383)	0.4227*** (0.0358)	0.4227*** (0.0358)	0.4717*** (0.04126)
<i>V</i>	0.3974*** (0.0161)	0.3861*** (0.0077)	0.4314*** (0.04142)	0.0884*** (0.0025)	0.1518*** (0.0052)	0.0543** (0.02600)	0.4267*** (0.0131)	0.4521*** (0.0131)	0.4897*** (0.04156)
<i>H</i>	0.0125* (0.0073)	0.0000 -	0.0005 (0.00047)	0.0402*** (0.0099)	0.1018*** (0.0220)	0.0049*** (0.00183)	-0.0102 (0.0251)	0.0122 (0.0302)	0.0054*** (0.00196)
<i>R</i>	0.0002*** (0.00003)	0.0015*** (0.0001)	0.0010*** (0.00009)	0.0136*** (0.0044)	0.0155*** (0.0047)	0.0114*** (0.00333)	0.0142*** (0.0045)	0.0171*** (0.0048)	0.0127*** (0.00342)
<i>(V/RE)%</i>	103.31%	100.38%	100.35%	255.86%	439.46%	142.89%	100.95%	106.95%	103.82%
<i>(H/RE)%</i>	3.26%	0.00%	0.12%	116.40%	294.64%	12.89%	-2.41%	2.89%	1.14%
<i>(R/RE)%</i>	0.06%	0.38%	0.23%	39.46%	44.82%	30.00%	3.36%	4.05%	2.69%
g	1.54%			0.74%			2.28%		

*, **, *** statistically significant at 10%, 5% and 1% levels respectively

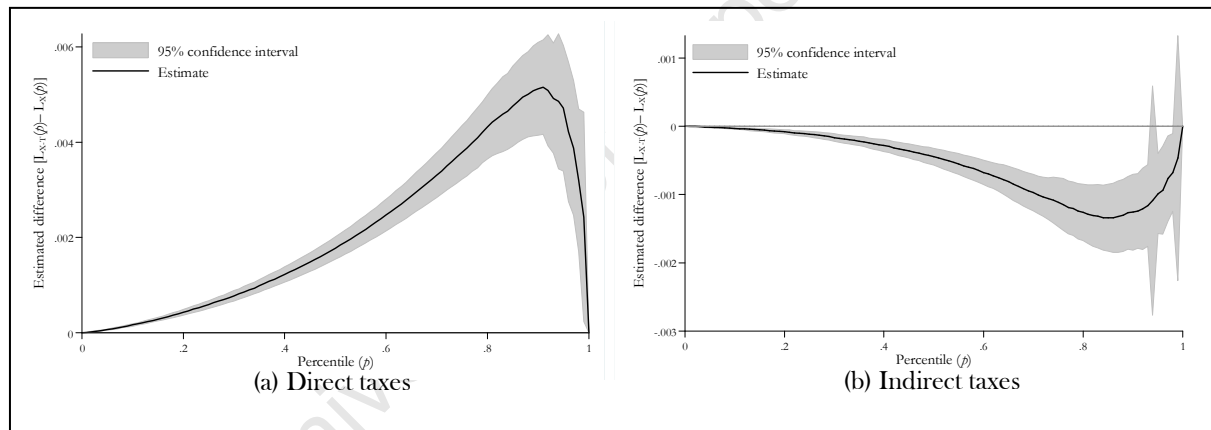
Standard errors in parenthesis

The values of *DJA* parameters are: $\varepsilon = 0.4$, $\nu = 1.5$

All figures are multiplied by 100 to enhance readability.

As shown in Table 7.3 the income redistributive effects of all direct taxes are positive irrespective of the choice of methods. The Gini based redistributive effect for total direct taxes are also presented in Figure 7.5(a). The effect is everywhere ($p \in [0,1]$) significantly positive. From Table 7.3 the redistributive effects of personal income tax and total direct taxes are statistically significant at 1% level across all models. This is only marginally significant for corporate income tax using the *AJL* and *UL* models and even not significant under the *DJA* approach. The vertical effect indicates that all direct taxes are significantly progressive because unequals were unequally treated in a favourable sense. These results conform to those presented earlier in Chapter 6. Horizontal inequity is significant for corporate income tax across all methods. This signifies the extent to which equals were not treated equally. As noted in the methodology section horizontal inequity reduces the extent of redistribution because when equals are not treated equally there is a possibility that the progressivity observed is dampened and could result in a loss in redistributive power of the tax.

Figure 7.5: *AJL* and *UL* decomposition of the redistributive effect of direct and indirect taxes



Personal income tax is horizontally equitable. This conclusion is because the horizontal effect is not statistically significant except for the *AJL* model where it is marginally significant at the 10% level. This means that equals were generally treated equally and it is a consequence of the structure of personal income tax rates. Total direct taxes apart from the *DJA* approach are horizontally equitable signifying that direct taxes treat equals equally. It is important to stress here that it is only in the *DJA* model that horizontal inequity maintains the classical definition. This is because there are no arbitrary groupings of individuals into close equal groups using income bands. Therefore horizontal equity in total direct taxes observed under the *AJL* and *UL* is a consequence of horizontal equity in the arbitrarily constructed bands. Note that $H^{AJL} \rightarrow 0$ and $H^{UL} \rightarrow 0$ as the size of the income bands are reduced considerably. This means

that based on the classical definition, the *DJA* approach shows that there is horizontal inequity in direct taxes. Total redistributive effects are generally reduced by significant re-ranking especially for corporate income tax. Re-ranking though significant for personal income tax is very small. As noted earlier, it should be borne in mind that while re-ranking is generally not expected for personal income tax, the results here use an income proxy in ranking households. Also personal income tax in South Africa is only based on natural persons and not the household as a unit. Because per adult equivalent incomes are used in this analysis it is possible that personal income tax could re-rank households due to household composition.

The relative contributions of vertical equity, horizontal inequity and reranking in personal income tax to changes in inequality as presented in Table 7.3 show that the vertical effect dominates for all methods. Under *AJL* personal income taxes would have been 3.3% more redistributive in the absence of differential treatment but this proportion is even lower under *UL* and *DJA*. The proportions represented by *UL* and *DJA* are very close. For example the *DJA* approach shows that personal income tax would have been 0.35% more redistributive in the absence of differential treatment while the *UL* approach states 0.38%. Under the *AJL* model, of the 3.31% effect due to differential treatment, most (3.26%) of this is attributed to horizontal inequity while only <0.1% is attributed to reranking. In contrast to this the *UL* and *DJA* attribute most of the differential treatments to reranking.

Corporate income tax across all methods redistributes income in favour of the poor. This is because of the reduction in inequality associated with payments. Even though this overall result is marginally significant, the vertical effect dominates. As shown in Table 7.3 *AJL* and *UL* show that there is more differential treatment in corporate income tax than that indicated by the *DJA* approach. The largest of these is obtained from the *UL* approach. Using the *UL* methodology, income redistribution would have increased by 340% in the absence of differential treatment. Again it is important to keep in mind that the results of H^{AJL} , H^{UL} and R^{AJL} are sensitive to the choice of income bands or intervals. This could have accounted for the larger influence of horizontal inequity over reranking under the *AJL* method. On average corporate income tax accounts for about 0.7% of household incomes, a proportion that is less than that accounted for by personal income tax.

As indicated in Table 7.3, total direct taxes used in financing health care account for about 2.3% of household income and this induces positive income redistribution across all the three methods. The reduction in inequality is a consequence of the positive income redistributive

effects of personal income and corporate income taxes. The reduction in inequality associated with direct taxes as a source of health care financing is mainly attributed to vertical equity effect. Under *UL* direct taxes would have been 6.95% more redistributive in the absence of differential treatment. About 4% of this differential treatment is caused by reranking while about 2.9% is caused by horizontal inequity. The horizontal effect for *AJL* is a false negative. In fact it is not statistically significant. This is because H^{AJL} was computed as a residual. Bilger (2008) had noted this in the case of health care financing in Switzerland and Urban and Lambert (2008) in the context of the *UL* methodology especially when the intervals are large.

Unlike direct taxes, indirect taxes generally induce significant negative income redistribution. The significant pro-rich redistribution was confirmed across all methods and for all the classes of indirect taxes contained in Table 7.4. The pro-rich redistribution associated with total indirect taxes using the decomposable Gini approach is generally consistent across all percentiles as shown in Figure 7.5(b). Post-payment inequality (G_{X-T}) was highest for value-added tax and then excise taxes and the least was for fuel levy. The implication of this is that the absolute value of the redistributive effect is higher for value-added tax than for excise taxes or fuel levy. This was consistent across all the models.

Table 7.4: Income redistributive effect of indirect taxes for health care financing, South Africa 2005/2006

	Value-added tax			Excise taxes			Fuel levy		
	<i>AJL</i>	<i>UL</i>	<i>DJA</i>	<i>AJL</i>	<i>UL</i>	<i>DJA</i>	<i>AJL</i>	<i>UL</i>	<i>DJA</i>
<i>RE</i>	-0.0522*** (0.0093)	-0.0522*** (0.0093)	-0.0576*** (0.01097)	-0.0384*** (0.0026)	-0.0384*** (0.0026)	-0.0425*** (0.00310)	-0.0181*** (0.0021)	-0.0181*** (0.0021)	-0.0204*** (0.00234)
<i>V</i>	0.0598*** (0.0035)	0.0909*** (0.0050)	-0.0571*** (0.01098)	-0.008*** (0.0011)	-0.0151*** (0.0014)	-0.0416*** (0.00316)	-0.0036*** (0.0007)	-0.0040*** (0.0001)	-0.0202*** (0.00234)
<i>H</i>	0.1119*** (0.0086)	0.1424*** (0.0127)	0.0000 (0.00000)	0.0301*** (0.0024)	0.0226*** (0.0023)	0.0005*** (0.00016)	0.0145*** (0.0013)	0.0141*** (0.0021)	0.0001 (0.00007)
<i>R</i>	0.0001*** (0.00001)	0.0008*** (0.00005)	0.0005*** (0.00003)	0.0003*** (0.00004)	0.0007*** (0.0001)	0.0004*** (0.00004)	0.00001*** (0.000001)	0.0001* (0.000006)	0.0000 (0.00000)
<i>(V/RE)%</i>	-114.48%	-174.06%	99.13%	20.94%	39.29%	97.88%	19.82%	22.00%	99.02%
<i>(H/RE)%</i>	-214.24%	-272.55%	0.00%	-78.31%	-58.75%	-1.18%	-80.13%	-77.56%	-0.49%
<i>(R/RE)%</i>	-0.24%	-1.51%	-0.87%	-0.76%	-1.94%	-0.94%	-0.04%	-0.43%	-0.48%
g	1.32%			0.20%			0.24%		

*, **, *** statistically significant at 10%, 5% and 1% levels respectively

Standard errors in parenthesis

The values of *DJA* parameters are: $\varepsilon = 0.4$, $\nu = 1.5$

All figures are multiplied by 100 to enhance readability.

Contrary to expectations, however, *AJL* and *UL* vertical effects for value-added tax were positive and significantly different from zero. It is important to note here that the original Kakwani index for value-added tax (π_K) was estimated at -0.0374 which showed that value-added tax is significantly regressive. However the Kakwani index ($\pi_{\bar{K}}$) used in the computation of the vertical effect in the case of *AJL* is slightly different. Based on (7.5) it is an index that is obtained in the absence of horizontal inequity in health care financing by grouping near-equals. All households in the group are given their group specific averages in the computation of $\pi_{\bar{K}}$ as opposed to π_K that uses observed household specific values. In the process of giving all households in each predefined “near-equals” their mean payments, the variability across households is lost and a ‘smoothened’ distribution is obtained. A similar procedure was used for the *UL*. Under the *UL* the payment rate specific to each group of near equals is deducted from their individual pre-payment incomes (Urban and Lambert, 2008). This also smooths the variability in the original distribution. In fact for the *AJL*, $V^{AJL} = (g / (1 - g)) \pi_{\bar{K}}$ such that the only factor that determines the sign of V^{AJL} is $\pi_{\bar{K}}$. This is because by definition $0 < g < 1$. When the groups were greatly increased (i.e. analogous to reducing the income intervals), V^{AJL} and V^{UL} become negative and most of the effects are transferred unto R^{AJL} and H^{UL} respectively. Unlike the *AJL* and *UL* models that require grouping of near-equals, *DJA* consistently maintained the significant negative relationship that was implied by the original Kakwani index of -0.0374. The vertical effect (V^{DJA}) as shown in Table 7.4 was negative and statistically different from zero. Horizontal (H^{DJA}) and reranking (R^{DJA}) effects were negligible.

Excise taxes used in financing health care that account for about 0.2% of total household income are consistently regressive and statistically significant across all the methods. Though horizontal inequity and reranking exist under the *AJL*, *UL* and the *DJA* approaches, their relative contributions to the overall income redistributive effect varies. Results from both the *AJL* and *UL* show that horizontal inequity dominates the vertical and reranking effects. For the *AJL* and *UL* approaches, horizontal inequity respectively accounts for about 78% and 59% of total differential treatments. The *DJA* result attributes most of the effects to vertical inequity (98%). The results for fuel levy, for all methods, also follow those for excise taxes.

Total indirect taxes, which account for about 1.8% of household income, generally redistribute income in favour of the rich. The results of the *DJA* as shown in Table 7.5 indicate that the

pro-rich redistribution originates from a regressive vertical effect. However the vertical effects for *AJL* and *UL* are positive and counter to *a priori* expectation. Just like value-added tax this is explained by the choice of income bands. Because of this, the majority of the differential effects are labelled as horizontal inequity. This comes back again to show that unlike *DJA*, *AJL* and *UL* do not measure classical horizontal inequity.

Table 7.5: Income redistributive effect of indirect taxes for health care financing, South Africa 2005/2006

	Indirect taxes		
	<i>AJL</i>	<i>UL</i>	<i>DJA</i>
<i>RE</i>	-0.1101*** (0.0120)	-0.1101*** (0.0120)	-0.1220*** (0.01430)
<i>V</i>	0.0493*** (0.0048)	0.0720*** (0.0045)	-0.1202*** (0.01443)
<i>H</i>	0.1588*** (0.0115)	0.1802*** (0.0143)	0.0007* (0.00042)
<i>R</i>	0.0006*** (0.0001)	0.0019*** (0.0001)	0.0011*** (0.00008)
$(V/RE)\%$	-44.75%	-65.38%	98.52%
$(H/RE)\%$	-144.24%	-163.70%	-0.57%
$(R/RE)\%$	-0.52%	-1.68%	-0.90%
g	1.76%		

*, **, *** statistically significant at 10%, 5% and 1% levels respectively

The values of *DJA* parameters are: $\varepsilon = 0.4$, $\nu = 1.5$

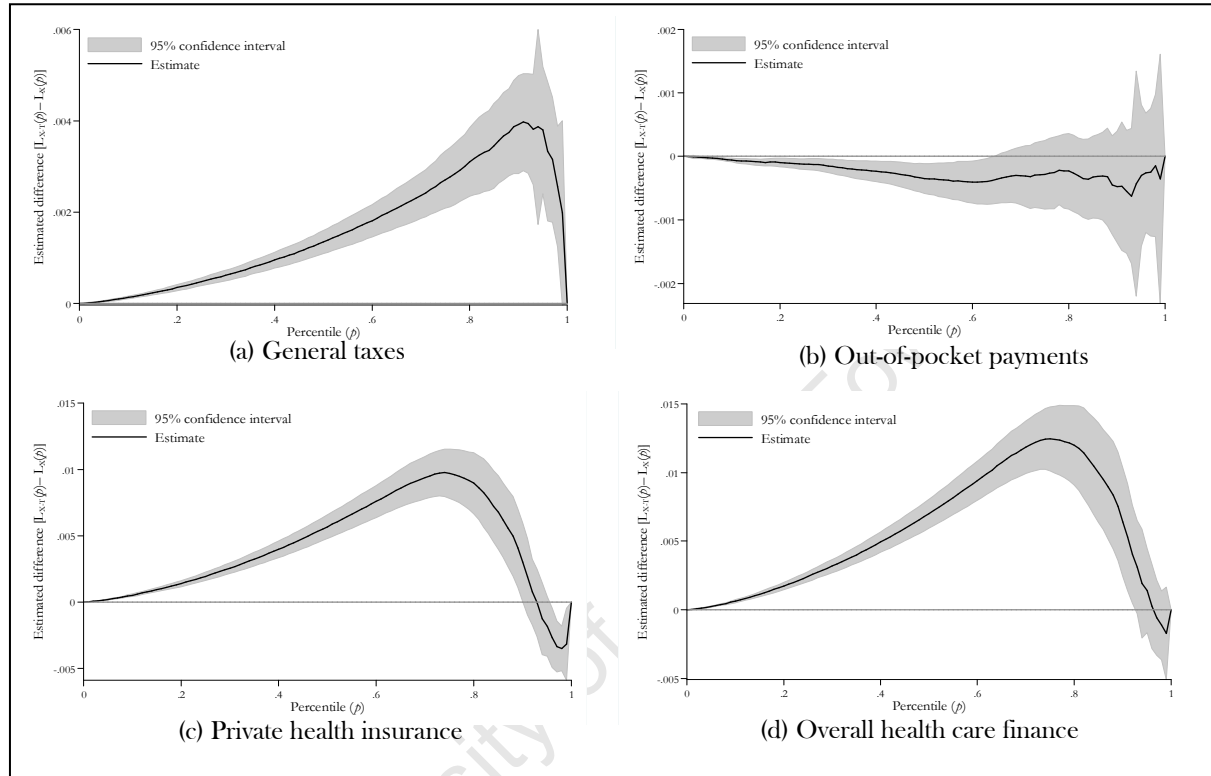
Standard errors in parenthesis

All figures are multiplied by 100 to enhance readability.

The comparative results presented in Table 7.6 for general taxes, private health insurance and out-of-pocket payments, show similar trends across all models. General taxes and private health insurance induce statistically significant positive income redistribution while out-of-pocket payments induce pro-rich redistribution. The graphic results of the Gini-based decomposition analysis are also shown in Figure 7.6. All models show that income redistribution associated with out-of-pocket payment is only mildly statistically significant (see Figure 7.6(b)). The vertical effects for general taxes and private health insurance follow positive *a priori* expectations. However the vertical effect for out-of-pocket payments under the *AJL* and *UL* models, though statistically significant, do not conform to expectation. These results indicate that out-of-pocket payments are statistically progressive. However the *DJA* results conform to original expectation that out-of-pocket expenditures are statistically proportional. Therefore the *AJL* and *UL* models fail to accurately predict the vertical component. The ‘false positive’ vertical effect for out-of-pocket payments under the *AJL* and *UL* models obviously

increased the horizontal effect such that this dominates all other effects. It is important to note that under the UL model for instance, total reranking (R^{UL}) and overall income redistributive effect are independent of the choice of close equal groups. Therefore the impact of choosing different income bands are accommodated by changes in V^{UL} and H^{UL} .

Figure 7.6: AJL and UL decomposition of the redistributive effect of health care financing



The relative contributions of the vertical, horizontal and reranking effects for both general taxes and private health insurance as contained in Table 7.6 further indicate that the vertical effects dominate for all models. However the differences that exist relate to how differential treatments are split between horizontal and reranking effects. AJL and UL models that do not measure classical horizontal inequity generally show that their horizontal effects dominate the reranking effects while the results from DJA model indicate otherwise. It should be recalled that under the AJL as noted earlier the relative share of horizontal effect and reranking can be changed simply by changing the size of the income bands. Similarly for UL because reranking is ‘fixed’, changing the income bands only changes the relative shares of reranking and horizontal inequity through changes in the vertical and horizontal effects. Reranking effects under the DJA model can however be reduced by decreasing the aversion to inequality parameter (ν). This effect will however be discussed extensively in later sections.

Table 7.6: Income redistributive effect of general taxes, private health insurance and out-of-pocket payments for health care financing, South Africa 2005/2006

	General taxes			Private health insurance			Out-of-pocket payments		
	<i>AJL</i>	<i>UL</i>	<i>DJA</i>	<i>AJL</i>	<i>UL</i>	<i>DJA</i>	<i>AJL</i>	<i>UL</i>	<i>DJA</i>
<i>RE</i>	0.3169*** (0.0329)	0.3169*** (0.0329)	0.3546*** (0.03687)	0.8241*** (0.1090)	0.8241*** (0.1090)	0.9401*** (0.12467)	-0.0482* (0.0285)	-0.0482* (0.0285)	-0.0574** (0.02889)
<i>V</i>	0.4283*** (0.0135)	0.5385*** (0.0179)	0.3759*** (0.03729)	1.1709*** (0.0786)	1.1512*** (0.0494)	1.1091*** (0.12942)	0.0269*** (0.0054)	0.0291*** (0.0036)	-0.0297 (0.03032)
<i>H</i>	0.0953*** (0.0290)	0.2015*** (0.0295)	0.0066*** (0.00201)	0.2114*** (0.0324)	0.1667** (0.0776)	0.0694*** (0.00729)	0.0572** (0.0264)	0.0534* (0.0293)	0.0124*** (0.00199)
<i>R</i>	0.0161*** (0.0049)	0.0202*** (0.0052)	0.0148*** (0.00375)	0.1354*** (0.0143)	0.1605*** (0.0158)	0.0996*** (0.00997)	0.0179*** (0.0041)	0.0238*** (0.0048)	0.0153*** (0.00320)
<i>(V/RE)%</i>	135.15%	169.95%	106.01%	142.08%	139.70%	117.98%	-55.89%	-60.31%	51.74%
<i>(H/RE)%</i>	30.09%	63.58%	1.86%	25.65%	20.23%	7.38%	-118.70%	-110.85%	-21.60%
<i>(R/RE)%</i>	5.07%	6.37%	4.17%	16.43%	19.47%	10.59%	-37.18%	-49.46%	-26.66%
g	4.04%			5.47%			1.63%		

*, **, *** statistically significant at 10%, 5% and 1% levels respectively

The values of *DJA* parameters are: $\varepsilon = 0.4$, $\nu = 1.5$

All figures are multiplied by 100 to enhance readability.

Finally as shown in Table 7.7 all the models indicate that overall health care financing induces positive income redistribution in favour of the poor (see also Figure 7.6(d) for the case of the decomposable Gini index). Similarly the vertical effect is consistently positive and statistically different from zero. This indicates that overall health care financing is progressive and places a lesser burden on the poor than on the rich. However the differences in the results of the three methods lie in the way the shares of differential treatments are allocated between horizontal inequity and reranking. Under the *AJL* overall health care financing would have been 48% more redistributive in the absence of differential treatment with horizontal inequity accounting for about 28% of this differential treatment. Similarly the *UL* model predicts that overall health care financing would have reduced inequality by 65% in the absence of differential treatment and the bulk of this proportion (42%) is also attributed to horizontal inequity. However the *DJA* model predicts that overall health care financing in South Africa would have been 22% more redistributive in the absence of differential treatment but the bulk of this (13%) is attributed to reranking while about 8% is due to horizontal inequity. Again the argument introduced for the case of general taxes and private health insurance in relation to choice of close equals seems to also hold here. Because the *DJA* model assesses classical horizontal inequity as unequal treatment of equals without creating arbitrary income bands its results were different and this is one of the major advantages of the *DJA* model over the *AJL* and *UL* models. However as pointed out by Duclos *et al.* (2003), the relative shares of reranking and horizontal inequity in their model could also be sensitive to the choice of the parameters ν and ε . In particular for any choice of ν the share of horizontal inequity in total differential treatment increases with ε . This is clearly so because H^{DJA} is more sensitive to changes in ε than to changes in ν . Similarly R^{DJA} is more sensitive to changes in ν than to changes in ε . Basically the ratio H^{DJA} / R^{DJA} increases with ε / ν (Duclos *et al.*, 2003). Because of this possibility sensitivity analyses were performed and the results of the sensitivity analyses show that the *DJA* results are robust to choice of the parameters within ‘reasonable’ limits.

Table 7.7: Income redistributive effect of overall health care financing, South Africa 2005/2006

Overall health care finance			
	<i>AJL</i>	<i>UL</i>	<i>DJA</i>
<i>RE</i>	1.1261*** (0.1278)	1.1261*** (0.1278)	1.2738*** (0.13699)
<i>V</i>	1.6653*** (0.0824)	1.8581*** (0.0432)	1.5499*** (0.15421)
<i>H</i>	0.3168*** (0.0643)	0.4736*** (0.1100)	0.1076*** (0.00941)
<i>R</i>	0.2224*** (0.0223)	0.2585*** (0.0241)	0.1685*** (0.01992)
<i>(V/RE)%</i>	147.88%	165.01%	121.68%
<i>(H/RE)%</i>	28.13%	42.05%	8.45%
<i>(R/RE)%</i>	19.75%	22.95%	13.23%
g	11.15%		

*, **, *** statistically significant at 10%, 5% and 1% levels respectively

The values of *DJA* parameters are: $\varepsilon = 0.4$, $\nu = 1.5$

All figures are multiplied by 100 to enhance readability.

The full decomposition of R^{UL} into R^{BG} , R^{WG} and R^{EG} are presented in Table 7.8. Generally, apart from personal income tax, between-group reranking accounts for the bulk of reranking for all progressive financing sources including total health care financing. This was also the case for out-of-pocket payments. Conversely for all indirect taxes, within group reranking dominates. For instance within group reranking accounts for about 84% of total reranking for value-added tax, 60.5% for excise taxes, 90% for fuel levy and 69% for total indirect taxes. Entire group reranking would require that an entire group out-ranks another in the redistributive process. As shown in Table 7.8 entire group reranking was estimated at zero for all the financing sources. This is a consequence of the nature of arbitrary grouping of households.

Table 7.8: *UL* decomposition of the re-ranking effects into between group, within group, and entire group re-ranking

Finance source	R	R(BG)	R(WG)	R(EG)	(R(BG)/R)%	(R(WG)/R)%
Personal income tax	0.0015*** (0.0001)	0.00022*** (0.00003)	0.00124*** (0.00011)	-	15.00%	85.00%
Corporate income tax	0.0155*** (0.0047)	0.01363*** (0.00443)	0.00185*** (0.00056)	-	88.05%	11.95%
Direct taxes	0.0171*** (0.0048)	0.01421*** (0.00453)	0.00293*** (0.00061)	-	82.91%	17.09%
Value-added tax	0.0008*** (0.00005)	0.00012*** (0.00001)	0.00066*** (0.00004)	-	15.25%	83.86%
Excise taxes	0.0007*** (0.0001)	0.00029*** (0.00004)	0.00045*** (0.00004)	-	38.98%	60.48%
Fuel levy	0.0001* (0.000006)	0.00001*** (0.0000)	0.00007*** (0.00001)	-	12.82%	89.74%
Indirect taxes	0.0019*** (0.0001)	0.00057*** (0.00006)	0.00128*** (0.00008)	-	30.81%	69.19%
General taxes	0.0202*** (0.0052)	0.01605*** (0.02024)	0.00415*** (0.00060)	-	79.46%	20.54%
Private health insurance	0.1605*** (0.0158)	0.13539*** (0.01431)	0.02506*** (0.00215)	-	84.38%	15.62%
Out-of-pocket payments	0.0238*** (0.0048)	0.01792*** (0.00411)	0.00592*** (0.00083)	-	75.17%	24.83%
Overall health care payments	0.2585*** (-0.0241)	0.22244*** (0.02228)	0.03603*** (0.00290)	-	86.06%	13.94%

*, **, *** statistically significant at 10%, 5% and 1% levels respectively
All figures are multiplied by 100 to enhance readability.

7.8 SENSITIVITY ANALYSES

This section presents the results of sensitivity analyses performed to assess the extent to which the results remain robust to changes in certain parameters. First, the robustness of the *AJL* and *UL* results was assessed to changes in the parameters of aversion to inequality. After this the robustness of the *DJA* results was assessed to changes in the parameters of aversion to inequality and uncertainty.

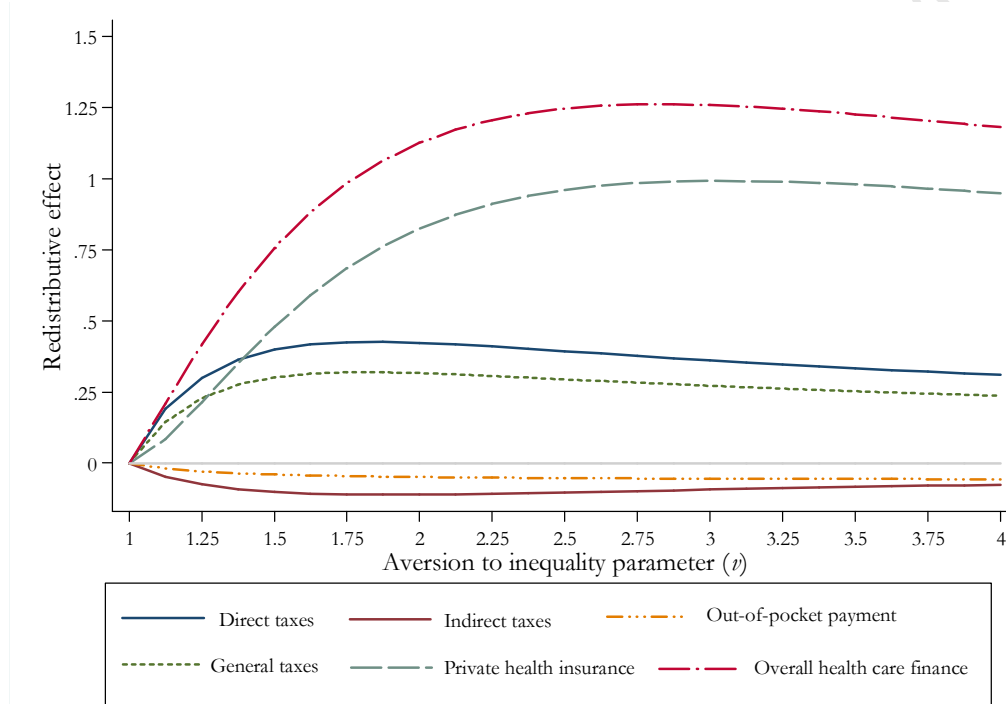
As introduced in Chapter 6 the Gini index is constructed based on an underlying value judgement. The standard Gini index equally weights all deviations from the line of equality ($p - L(p)$) by a constant scalar of 2. However the society may be interested in weighting more, the deviations of the poor, relative to those of the rich. In order to accommodate this, the single-parameter Gini index formulation was used in assessing redistribution as:

$$G_X(\nu) - G_{X-T}(\nu) = \int_0^1 (L_{X-T}(p) - L_X(p)) \kappa(p; \nu) dp \text{ for } \nu \geq 1 \quad (7.37)$$

where $\kappa(p; \nu) = \nu(\nu - 1)(1 - p)^{(\nu-2)}$ is the single parameter weighting function and ν is the parameter of aversion to inequality. This is the relative weights that the society or policy maker places on inequalities occurring among the poor relative to those among the rich (Yaari, 1988).

The results from using (7.37) are presented in Figure 7.7 for direct taxes, indirect taxes, general taxes, out-of-pocket payments, private health insurance and overall health care financing. Here different values of ν were used in assessing the sensitivity of the results.

Figure 7.7: Sensitivity of the redistributive effect to choice of inequality aversion parameters under the decomposable Gini index



Note: the redistributive effects have been scaled up by 100

As shown in Figure 7.7 especially for values of $\nu > 2$ the results of income redistribution for all health care financing sources remain robust to any choice of value for the parameter. Consistently income redistribution associated with private health insurance dominates that from direct taxes and general taxes. Similarly the negative income redistribution associated with out-of-pocket payments dominates the negative income redistribution resulting from indirect taxes.

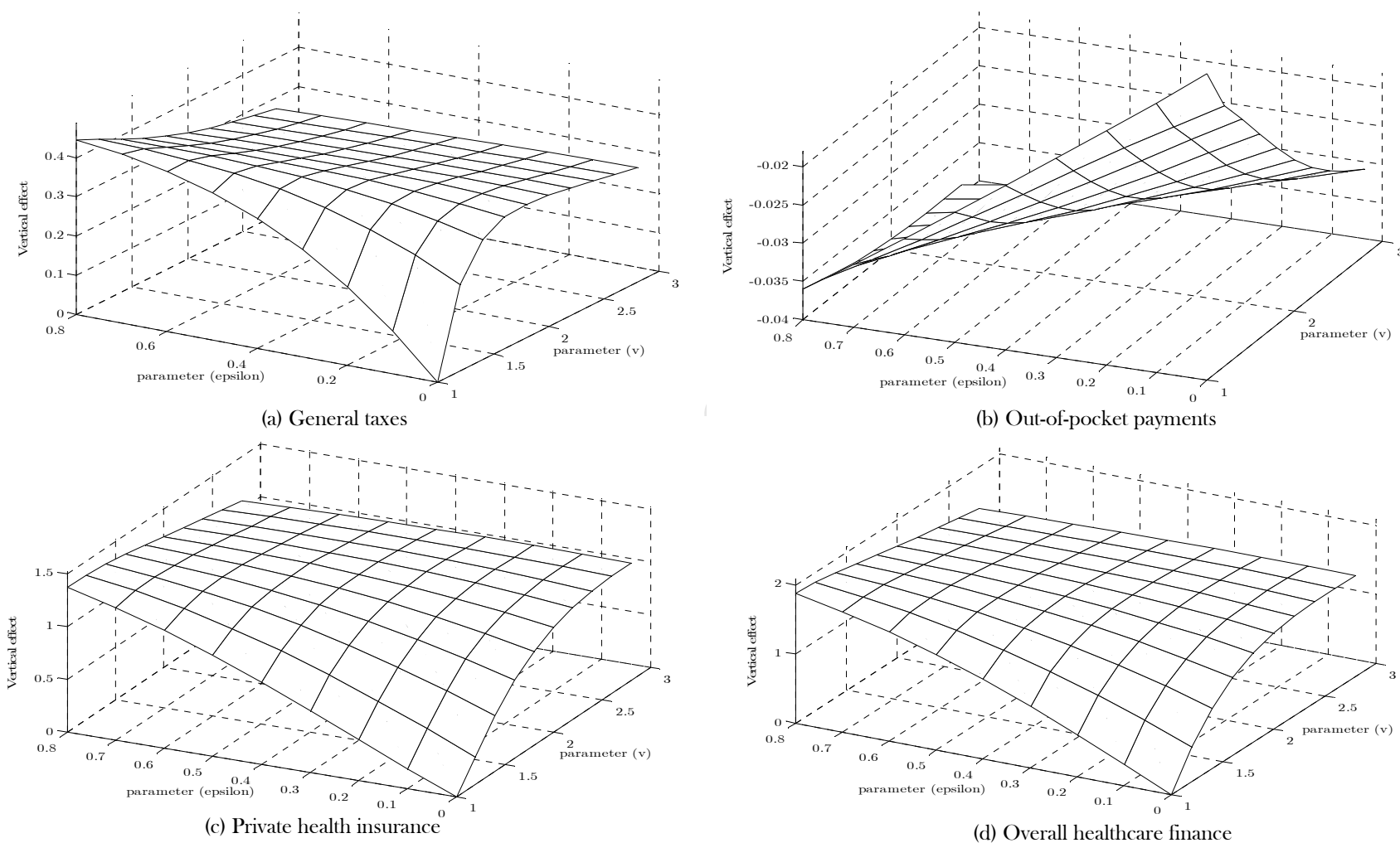
Because income redistribution in overall health care financing is a weighted sum of income redistribution generated by all the other financing sources, it is not surprising that its curve

dominates all other curves. In fact this is particularly the case where negative redistribution is very small. For values of $\nu < 1.4$, there are points where the distributions of direct taxes and general taxes dominate the distribution of private health insurance. However values of $\nu < 2$ represent cases where more weights are attached to deviations $(L_{x-T}(p) - L_x(p))$ occurring among the rich than those occurring among the poor. This may not be particularly appealing given that we are interested in accounting for relative deprivations occurring more among the poor (Yaari, 1988). Therefore the consistent ordering obtained when $\nu > 2$ indicates the robustness of income redistribution in South Africa to choice of parameter of aversion to inequality. The results obtained when $\nu = 2$ are the same as those presented earlier because $\kappa(p; \nu) = 2$.

Turning now to the *DJA* model the results of the sensitivity of RE^{DJA} , V^{DJA} , H^{DJA} and R^{DJA} to changes in the parameters ν and ε have been shown. As shown in Figure 7.8 apart from extreme values of ν and ε the results are generally robust to the choice of parameter values. This is the same for general taxes, private health insurance, out-of-pocket payments and overall health care financing. The value of V^{DJA} only changes marginally when the values of the parameters are changed. Similarly as shown in Figure 7.9 the value of H^{DJA} is robust to changes in values of the parameters within reasonable limits. It is only at extreme values such as $\varepsilon \rightarrow 0$ and $\nu \rightarrow 1$ that the values of H^{DJA} become very responsive. H^{DJA} is however more responsive to ε than to ν . Reranking effects unlike the horizontal effects indicate that R^{DJA} is more responsive to ν than to ε . As shown in Figure 7.10 apart from the cases when $\nu < 1.5$, the results are generally robust to the choice of the parameters. As noted earlier the relative shares of horizontal inequity and reranking under the *DJA* model could be sensitive to the choice of parameters. Specifically H^{DJA} increases with ε and R^{DJA} increases with ν such that H/R increases with ε/ν . However based on the results of the sensitivity analyses for both H^{DJA} and R^{DJA} it is clear that within reasonable limits (i.e. $\nu \geq 1.5$ and $\varepsilon \geq 0.2$) the ratio H/R would remain robust.

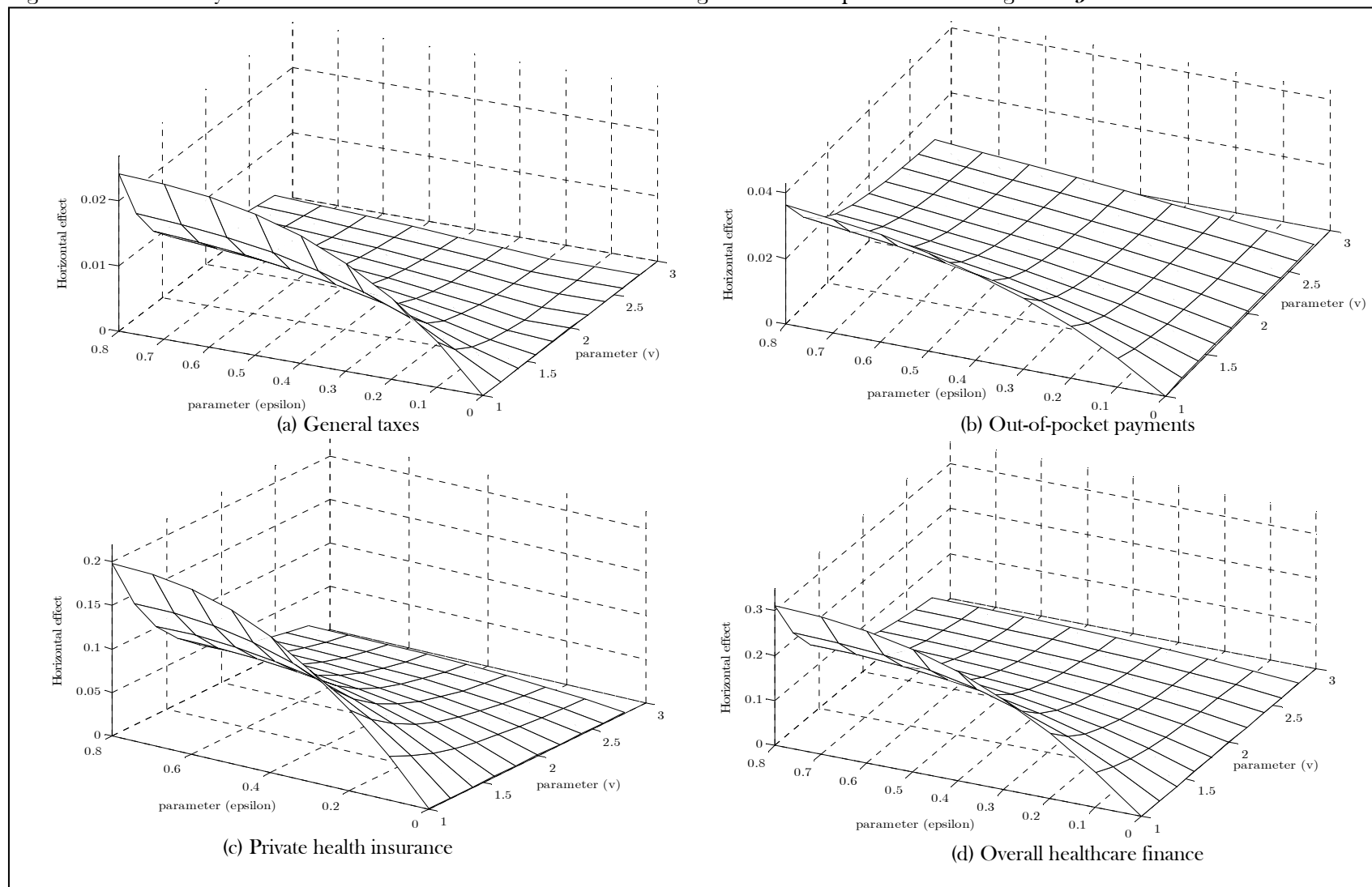
The sensitivity results for overall income redistributive effect (RE^{DJA}) shown in Figure 7.11 indicate also the robustness of the results to choice of the parameters. This is particularly the case for $\nu \geq 1.5$ and $\varepsilon \geq 0.2$. Therefore the results shown in Figures 7.8 - 7.11 show the robustness of the results of the *DJA* model to choice of values for the parameters ν and ε .

Figure 7.8: Sensitivity of the vertical effect of health care financing to choice of parameters using the *DJA* model



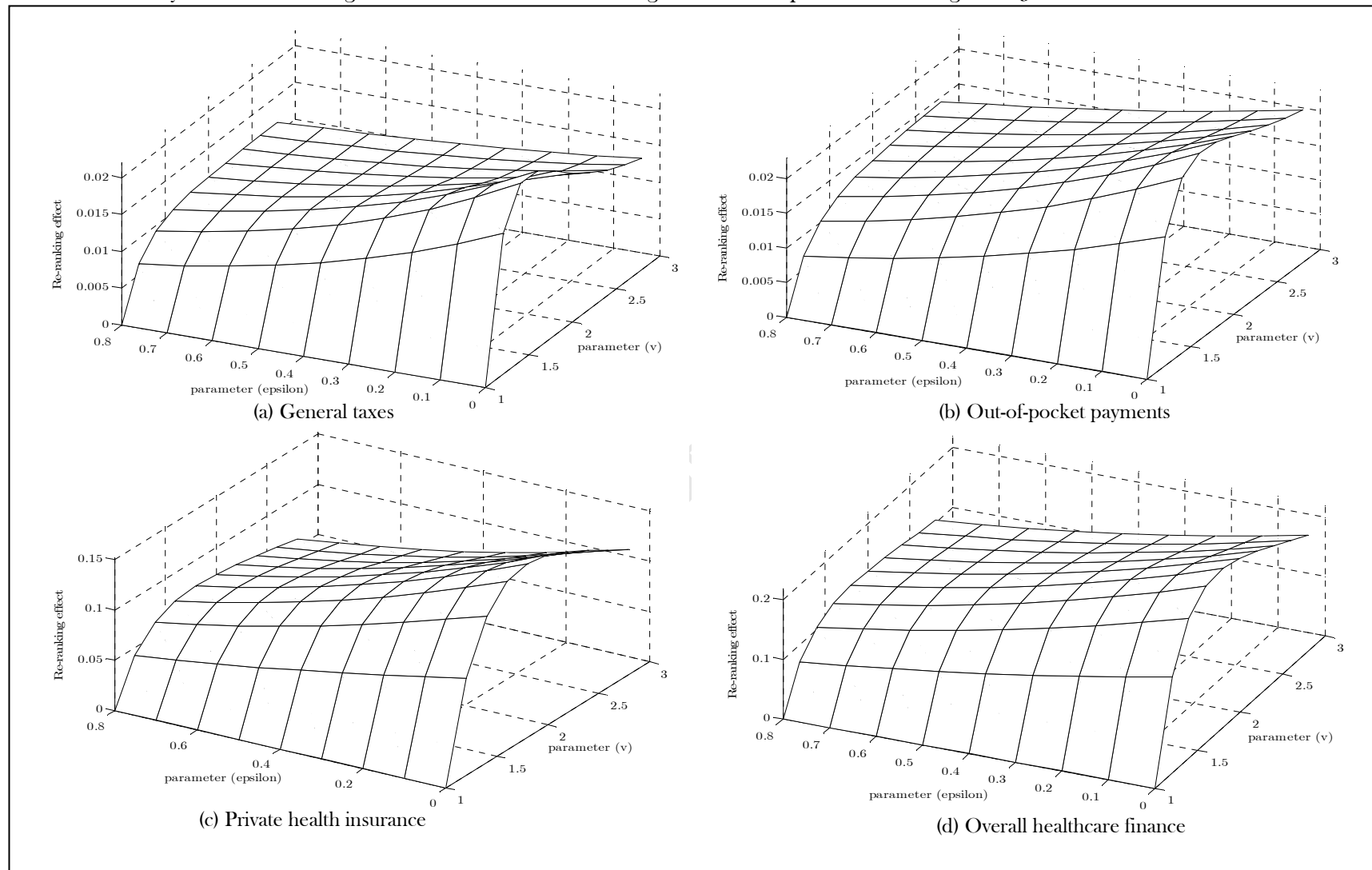
Note: the vertical effects have been scaled up by 100

Figure 7.9: Sensitivity of the horizontal effect of health care financing to choice of parameters using the *DJA* model



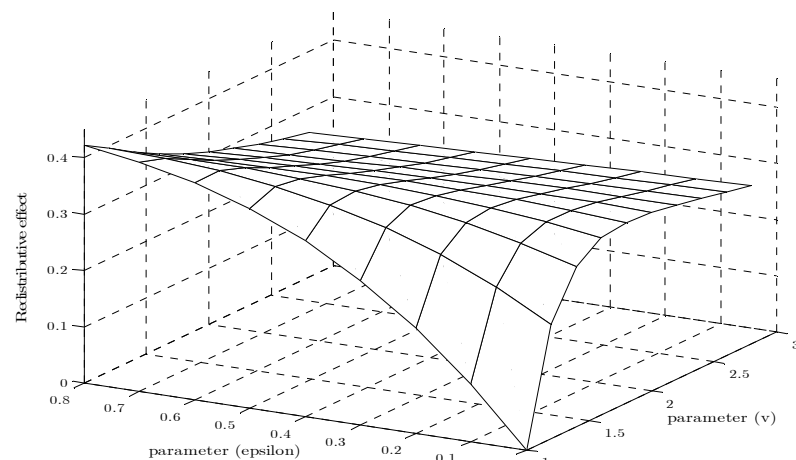
Note: the horizontal effects have been scaled up by 100

Figure 7.10: Sensitivity of the re-ranking effect of health care financing to choice of parameters using the *DJA* model

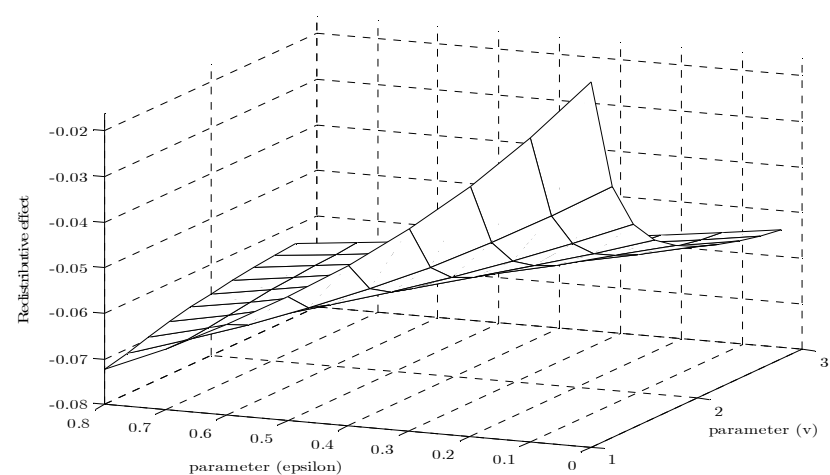


Note: the reranking effects have been scaled up by 100

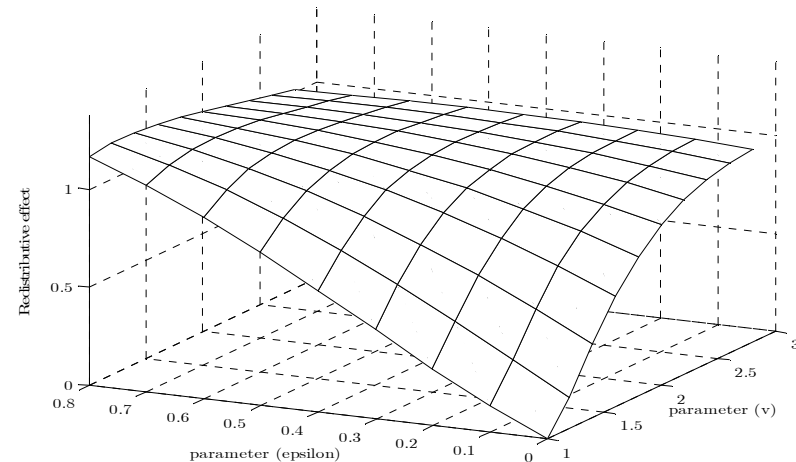
Figure 7.11: Sensitivity of the redistributive effect of health care financing to choice of parameters using the *DJA* model



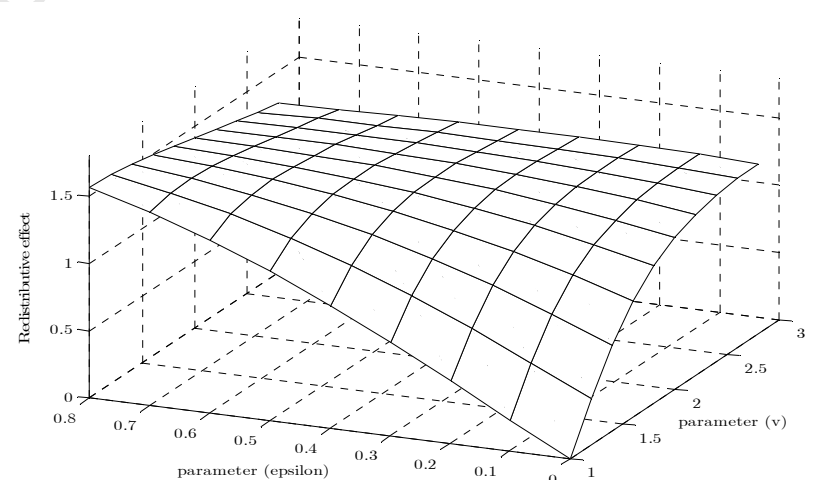
(a) General taxes



(b) Out-of-pocket payments



(c) Private health insurance



(d) Overall healthcare finance

Note: the redistributive effects have been scaled up by 100

7.9 DISCUSSION

Health care financing, whether through general taxes, health insurance contributions or out-of-pocket payments, involves some intra- and inter- household decisions that are likely to impact on the distribution and redistribution of income across households. Though the main purpose of health care financing is not to redistribute income (Cavagnero and Bilger, 2010), the extent to which income distribution is affected by health care financing is an important policy issue. It is determined by the way individuals and households end up being treated by the health care financing system. When the system is progressively financed in such a way that it imposes less financial burden on the poor relative to the rich it usually redistributes income away from the rich to the poor. Minimising differential treatments in the form of horizontal inequity and reranking of households also improves the extent of income redistribution associated with health care financing. Different countries have different health care financing mixes. Some countries such as Denmark, Ireland, Niue Island, Swaziland, Sweden, Thailand and United Kingdom rely more on general taxes while some countries, mainly in the developing world such as Cameroon, Côte d'Ivoire, Nigeria, Pakistan, Sierra Leone and Zambia rely heavily on direct out-of-pocket payments. The relative progressivity and shares of these sources to overall health care financing to a large extent determine the degree to which health care financing can redistribute income. This does not necessarily depend on whether the country in question is developed or is still developing. In fact in some developed countries such as Denmark, United States, Switzerland, Portugal, Germany and the Netherlands, overall health care financing has been shown to increase income inequality (Wagstaff and van Doorslaer, 1997, van Doorslaer *et al.*, 1999, Bilger, 2008). This unfavourable increase in income inequality translates into a negative redistribution of income away from the poor to the rich.

Equity is extolled as a key objective in health policy documents in all countries of the world. The major issue in relation to pursuing equity is to avoid placing an increasing burden on the poor who are already overburdened by disease, inequality and poverty (Wagstaff, 2002a). However some countries still struggle to strike a balance between competing challenges in the health sector and have not been able to guarantee an equitable distribution of health care financing burden and also access to health care for the majority of its citizens.

While there are competing methodologies that have been proposed, mainly from the tax literature, to understand the extent of redistribution associated with some tax regime or

financing system, their application to studying health care financing have been limited. It is only a few studies, mainly within developed countries that have provided empirical evidence on the extent to which health care payments could lead to differential treatment of individuals and households. These issues should however not be limited to developed countries as they are also central in developing countries that are striving to achieve the United Nations Millennium Development Goals (MDGs). The dearth of studies that examine the impact of health care financing on income inequality in developing countries is largely attributed to data insufficiency. However reliable data are becoming increasingly available that researchers can draw upon to provide evidence in the context of developing countries where issues of poverty, unemployment, inequality and disease are endemic. In this regard the current study provides evidence in the context of South Africa – a developing country - to show the extent to which health care payments impact on household welfare.

The study using the *DJA* approach has shown that health care financing via indirect taxes increases income inequality in South Africa. This result is consistent with those reported in other studies. Using the *DJA* methodology it was shown that the redistributive effect associated with financing health care via indirect taxes is -0.0012 and in countries such as Argentina (in 2002), Italy (in 1991), Portugal (in 1990) and United Kingdom (in 1992) using either the *AJL* or *DJA* models, the income redistributive effect of health care financing via indirect taxes was estimated at -0.0011, -0.0014, -0.0012 and -0.0018 respectively (van Doorslaer *et al.*, 1999, Cavagnero and Bilger, 2010). In all previous studies the vertical effect is the major driver of the negative redistribution associated with indirect taxes. In some countries such as Denmark, Germany, Italy, Netherlands, Sweden and Switzerland the pro-rich vertical effect accounts for all of the negative income redistribution (Wagstaff and van Doorslaer, 1997, van Doorslaer *et al.*, 1999). Generally between 60% and 100% of the pro-rich redistribution is attributed to regressive vertical effect in indirect taxes. In South Africa the *DJA* approach estimates put this at 98.5%. Differential treatments that account for about 1.5% of the loss in redistributive power of indirect taxes used in financing health care in South Africa can be attributed largely to the non-discriminatory nature of indirect taxes. It could also be associated, to a lesser extent, with the consumption patterns that exist among households with similar incomes. In the case of Portugal for example differential treatment was associated with multiple value-added tax rates on different goods and services (van Doorslaer *et al.*, 1999).

Internationally direct taxes on the other hand have been shown to reduce inequality in income. In all countries where research has been conducted and reported, the redistributive effect was positive. The same is true for South Africa as this study confirms. The redistributive effect of direct taxes in South Africa using the *DJA* approach was estimated at 0.0047 with the vertical effect dominating. The positive redistribution usually associated with financing health care via direct taxes is mainly the result of progressive direct taxes. In Argentina, Denmark, Finland, Germany Ireland, Italy, Netherlands, Portugal, Sweden, Switzerland, the United Kingdom and the United States direct taxes were reported as progressive and the contribution of the vertical effect to overall redistribution ranged between 100% and 143% (Wagstaff and van Doorslaer, 1997, van Doorslaer *et al.*, 1999, Bilger, 2008, Cavagnero and Bilger, 2010). The reported differential treatment (horizontal inequity and reranking) in the case of Portugal is linked to differences in tax structure between wage earners and the self-employed (van Doorslaer *et al.*, 1999) such that individuals that earn a similar income end up making differential tax payments. In countries such as Denmark, Sweden and the United States differential treatments, though small, were attributed to geographical/regional differences in tax structure (van Doorslaer *et al.*, 1999). In some countries such as Germany, Italy and Ireland the pro-poor vertical effect accounts for all of the redistribution because there were no significant differential treatments (van Doorslaer *et al.*, 1999).

The combination of a pro-rich redistribution associated with indirect taxes and a pro-poor redistribution associated with direct taxes used in financing health care have produced a pro-poor redistributive effect for general taxes. In this study the redistributive effect of financing health care via general taxes in South Africa using the *DJA* approach was estimated at 0.0035 with the vertical effect accounting for over 106% of this effect. Internationally, albeit for developed countries where results are available, financing health care through general taxes also reduces income inequality. For example in Denmark (in 1987) using the *AJL* method the redistributive effect of general taxes was estimated at 0.0024 while in Switzerland (1992) it was estimated at 0.0035. In all these countries the pro-poor vertical effect associated with general taxes dominates the horizontal and reranking effects. In some countries such as Germany, Italy and the Netherlands there were no significant differential treatments associated with general taxes used to finance health care. This is because 100% of the redistributive effect is as a result of a progressive vertical effect. Generally most studies conclude that general taxes would have been between 2% and 18% more redistributive in the absence of differential treatment.

The result of the redistributive effect of private health insurance contributions in South Africa indicates that it reduces income inequality. The redistributive effect estimated using the *DJA* approach was 0.0094 with a significantly dominant progressive vertical effect. The reduction in inequality caused by private health insurance is related to the distribution of the enrollees. It is important to bear in mind here that unlike general taxes where every citizen is entitled to services provided at public facilities, it is only those who are privately insured that can access benefits associated with private health insurance (Ataguba and McIntyre, 2012). In fact the poor, because they do not purchase private health insurance, do not benefit from the contributions. In South Africa the progressive vertical effect dominates the horizontal inequity and reranking effect. The presence of differential treatment is largely attributed to individuals facing the same premium schedule irrespective of income level. However some individuals mitigate this by choosing lower cost and less generous service benefit options. This could have accounted for the smaller than expected differential treatments observed.

Elsewhere the redistributive effect associated with private health insurance can be negative or positive depending on the nature of health insurance in the country (i.e. how premiums are structured and paid for, whether the insurance is supplementary or not, whether it is only the rich who are covered, and whether enrolment is based on individuals or households). In the United States and formerly Switzerland (pre their 1996 reform) where private health insurance is dominant and not only restricted to the rich, pro-rich redistribution was reported while countries such as the United Kingdom, Portugal and Italy have pro-poor redistribution because it is only the rich that can afford supplementary cover (van Doorslaer *et al.*, 1999). The non-zero values associated with horizontal inequity in Portugal and the United Kingdom for instance were associated with likely differential treatments induced by those with similar incomes where some choose to purchase health insurance cover and others do not. In Switzerland insured individuals may choose more generous insurance packages to provide cover for luxurious hospital services (van Doorslaer *et al.*, 1999) thereby creating differentials that could be labelled as horizontal inequality rather than inequity. In countries such as Argentina where private health insurance accounts for about 32% of total private health expenditures, pro-poor income redistribution has been reported (Cavagnero and Bilger, 2010). The major driver of the pro-poor redistribution is the progressive nature of private health insurance as insurance premiums are too expensive for the poor to enrol. Cavagnero and Bilger (2010) attributed differential treatments in this case to heterogeneity in risk aversion and

differences in health status. Whether private health insurance induces pro-rich or pro-poor income redistribution, the vertical effect dominates the horizontal and reranking effects.

As with indirect taxes, direct out-of-pocket payments are associated with negative income redistribution. In this study the *DJA* model estimates the redistributive effect at -0.0006 with a fairly dominant (51%) regressive vertical effect. This result is similar to those obtained from Vietnam using the *AJL* approach and West Bank (Palestine) using the *UL* methodology. In Vietnam (in 1993) it was estimated that the regressive vertical effect accounted for about 47% of total income redistribution (Wagstaff and van Doorslaer, 2001) while in West Bank it accounted for about 43% (Abu-Zaineh *et al.*, 2009). In the case of South Africa, even though out-of-pocket expenditures make up about 14% of total health care financing, the pro-rich redistribution induced is not equitable and would reduce the extent to which overall health care financing redistributes income. The non-zero values of horizontal inequity and reranking is largely a result of the non-discriminatory nature of out-of-pocket payments and health service utilization patterns especially for the rich and insured that make co-payments. Large and significant differential treatment associated with out-of-pocket payment in South Africa can be described as inequitable because of the stochastic nature of illness. Some individuals with the same level of income could randomly fall sick and this will inherently create randomness in out-of-pocket payments. Co-payments by members of private health insurance that make up over 60% of out-of-pocket expenditures in South Africa (see McIntyre, 2010b) are a result of institutional failures in health insurance arrangements and this induces differential treatment of individuals with the same or similar level of income.

Internationally, even though there are cases where out-of-pocket payments result in substantial differential treatment, in general most studies show that the vertical effect substantially dominates the horizontal and reranking effects. For example the study in the Netherlands using the *AJL* method indicates that the vertical effect accounts for about 82% of the negative redistribution associated with out-of-pocket expenditures (Wagstaff and van Doorslaer, 1997) and in Switzerland using *DJA* methodology Bilger (2008) reports that the vertical effect accounts for about 78% of the negative redistribution resulting from out-of-pocket payments. Some studies have reported negative redistribution with out-of-pocket payments but with a progressive vertical effect (see Ichoku, 2006, Cavagnero and Bilger, 2010). In such countries the combination of horizontal and reranking effects dominates the progressive vertical effect. For instance in Argentina (in 1997) using the *DJA* approach about 182% of the reduction in

redistributive power of out-of-pocket payment is caused by differential treatment in the form of horizontal inequity and reranking (Cavagnero and Bilger, 2010) and this was even higher in the case of Nigeria (Ichoku, 2006). It is only in Argentina (in 2002) that out-of-pocket payments were recorded to be progressive and redistribute income favourably towards the poor (Cavagnero and Bilger, 2010). The general negative redistribution associated with out-of-pocket payments points to the inequitable nature of such payments. Even when they turn out to have a progressive vertical effect, it is a result of the exclusion faced by the poor on the grounds that they cannot afford the cost of treatment (Cavagnero and Bilger, 2010). In the context of developing countries where the bulk of health care financing is through out-of-pocket payments (Ataguba, 2012b) this is a major cause for concern. If this general trend of pro-rich redistribution is true for these countries, the extent of inequity in the overall health care system is likely to be huge.

Overall health care financing in South Africa was shown to induce a pro-poor redistribution of income. This reduction in income inequality is attributed to the pro-poor redistribution caused by general taxes and private health insurance which dominates the fairly pro-rich redistribution associated with out-of-pocket payments. The reduction in income inequality associated with total health care financing in South Africa using the *DJA* approach was 0.0127 with the progressive vertical effect accounting for over 121% of this effect. This is comparable to that reported elsewhere in Argentina ($RE = 0.018$) in 2002 using also the *DJA* framework (Bilger, 2008). Generally many countries record negative income redistribution for overall health care financing which is caused mainly by regressive vertical effects. These countries include Denmark, Germany, Netherlands, Palestine, Portugal, Switzerland and the United States (see Wagstaff and van Doorslaer, 1997, van Doorslaer *et al.*, 1999, Abu-Zaineh *et al.*, 2008, Bilger, 2008, Abu-Zaineh *et al.*, 2009). However a few countries such as Argentina, France, Italy and Sweden (see van Doorslaer *et al.*, 1999, Cavagnero and Bilger, 2010) have recorded a positive income redistributive effect with overall health care financing. Irrespective of the nature of redistribution caused by total health care financing, the progressive or regressive vertical effect generally dominates.

7.10 METHODOLOGICAL DEBATES

One major issue with assessing income redistribution associated with health care financing or taxes is the choice of close-equals. This is particularly the case for the decompositions based

on the Gini index (AJL and UL). The choice of close-equal groups could lead to perverse judgements, not with the overall redistribution but with the respective components of redistributive effects. For example H and R are by definition non-negative quantities. However reducing the bandwidths used in constructing close equal groups usually underestimates the H and could even produce negative values for H (Urban and Lambert, 2008). Under the AJL model different procedures have been proposed and used to compute V^{AJL} . One is based on computing the Kakwani index that assumes away horizontal inequity (cf. Wagstaff and van Doorslaer, 1997, van Doorslaer *et al.*, 1999, Wagstaff *et al.*, 1999a). Another method basically computes the vertical effect as the difference between the Gini index of pre-payment income and the between-group Gini index of post-payment income. Different values of V^{AJL} are obtained depending on the procedure adopted (see Urban and Lambert, 2008). Because the horizontal effect is usually computed as a residual, to ensure exact decomposition, different values of V^{AJL} would translate into different values of H^{AJL} . In fact as shown in the case of direct taxes, $H^{AJL} < 0$. This was to compensate for either a large value estimated for R^{AJL} or a smaller value estimated for V^{AJL} . Therefore the computation of V^{AJL} , H^{AJL} and R^{AJL} are affected by the choice of close-equal groups. Again the results presented for South Africa indicate a counter intuitive sign on the vertical effect (V^{AJL}) for value-added tax and total indirect taxes. This counter intuitive sign was mainly a result of the construction of close-equal groups and the averaging effect induced by the way the Kakwani index ($\pi_{\bar{K}}$) was computed (see also Urban and Lambert, 2008). Also R^{AJL} does not measure full reranking (Urban and Lambert, 2008). These pose challenges to the results of the AJL methodology.

The revised methodology as contained in the work of Urban and Lambert (2008) also did not deal with the problem of choosing close equals. In fact what their methodology does is to push the effect of the choice of close equal groups on to V^{UL} and H^{UL} and on the relative values of within-group (R_{WG}^{UL}), between-group (R_{BG}^{UL}) and entire group (R_{EG}^{UL}) reranking in total reranking. This is because by construct, total redistribution (RE) and total reranking (R^{UL}) are invariant to the choice of bandwidth. As also shown in the results of the decomposition, V^{UL} is counter-intuitive for value-added tax and total indirect taxes. This was again the result of the choice of bandwidth.

The challenge of choosing close equals have been critically examined by for example van de Ven *et al.* (2001). To get around the choice of bandwidths and close equal groups they outline an optimal procedure to follow in making the choice. They note that this optimum choice of

bandwidth is that which maximises the vertical effect. This works through some averaging and appropriation effects (see Chapter 4 for details on these effects). However this is only applicable to progressive financing as there is likely to be no maximum vertical effect for regressive financing (Bilger, 2008). The real issue here is that decisions about close equals should not be based on mathematical algorithms or equations that seek to maximise some abstract construct. Close equals or more strictly ‘equals’ are determined by people based on their own value judgements. This is what drives the literature of relative deprivations (Runciman, 1966, Runciman and Garrison, 1972) and feelings of resentment (see Duclos and Lambert, 2000). This is a value-laden issue that should not be reduced to mathematical derivation. The question should be: who judges who is equal to oneself? Because the answer to this question is normative, arbitrary groupings further compound our understanding of equals.

In the context of South Africa the arbitrary choice of bandwidth is fraught with challenges that originate from the wide disparities in income. As illustrated in Figure 7.4 the distribution of income is skewed and the density is concentrated within a relatively narrow interval. To get around this skewed distribution of income as shown in Figure 7.1 close equals were arbitrarily constructed by grouping households into 100 groups based on per adult equivalent income. This is consistent with the use of percentiles to describe the distribution of income. Each group is basically a percentile of the population. Though the grouping had no effect on the overall redistributive effect (RE), it however affects the results of the individual components of redistribution for the AJL and UL decompositions as discussed earlier. The DJA approach was then used in this context because by construct it does not require any arbitrary grouping of ‘close-equals’. The results from the DJA decomposition conform to *a priori* expectations based also on the results of the incidence presented in Chapter 6. The major challenge with the DJA approach is the issue of choosing values for the parameters of aversion to inequality and to uncertainty. However sensitivity analyses that were performed to examine the impact of these choices as shown in Figures 7.8 – 7.11 conclude that the results presented are robust to the choice of parameter values within reasonable limits. Because the DJA model does not impose arbitrary specification of close-equal groups it performed better and this justifies its relevance in the context of South Africa as presented and discussed in this chapter.

7.11 CONCLUSION

This chapter critically assesses the income redistributive effect associated with health care financing in South Africa. Financing health care through direct taxes, private health insurance and general taxes were shown to induce positive redistributive effects while indirect taxes and out-of-pocket payments induce negative income redistributive effects. The combined effect of these results is a pro-poor redistributive effect associated with total health care financing in South Africa. In all cases the vertical effect dominates the horizontal and reranking effects but this is less so for out-of-pocket payment and private health insurance where substantial differential treatments occur. Internationally there is no consistent pattern of income redistribution that is associated with total health care financing. However consensus has emerged in terms of direct and indirect taxes. Direct taxes induce pro-poor redistribution while indirect taxes induce negative redistribution. Similarly the result for private health insurance indicates that a pro-rich redistributive pattern is associated with schemes that tend to cover both the rich and the poor while pro-poor income redistribution is associated with schemes that cover only the rich who can afford payments. Out-of-pocket expenditures show mixed results ranging from a pro-poor to pro-rich redistributive effect. In the case of South Africa differential treatments arising from out-of-pocket payments and private health insurance were described as inequitable because of the nature of payments and enrolment of individuals. Sensitivity analyses were used to assess the impact of certain assumptions on overall redistribution and the conclusions were that the results are robust to changes in the values of the parameters within reasonable limits. The comparison of models in the context of South Africa, with high income inequality, also showed that the *DJA* approach was consistent and produced plausible results compared to the *AJL* and *UL* models that require arbitrary choice of income bandwidth.

Chapter 8

IMPLICATIONS OF THE INCIDENCE AND INCOME REDISTRIBUTIVE EFFECT FOR EQUITY IN HEALTH CARE FINANCING IN SOUTH AFRICA

8.1 INTRODUCTION

The implications of the incidence results and the decomposition analyses of health care financing presented and discussed in Chapters 6 and 7 are discussed in this chapter. The preceding chapters have not only shown that overall health care financing in South Africa is progressive but that it could also lead to differential treatments in the forms of reranking and horizontal inequity. This chapter thus attempts to answer the questions: what are the options in South Africa to improve the progressivity and income redistribution associated with health care financing? How can we make the South African health care financing system as equitable as possible? The discussions here are also rooted in the ability to pay principle.

This chapter also discusses the contribution of this thesis to the literature on health care financing. It outlines the limitations of the study and some recommendations for future research. The chapter also provides an overall recommendation for ensuring an equitable health care financing system in South Africa.

8.2 IMPLICATIONS FOR EQUITABLE HEALTH CARE FINANCING IN SOUTH AFRICA

It is the case that most societies would prefer a progressive health care financing system over a regressive one (see for example Wagstaff, 2002b). This is because it promotes equity in health care financing and is built on some sort of social solidarity. In South Africa there is a large gap between the haves and the have-nots as reflected in the distribution of income. The richest 10% of the population accounts for about 51% of total income compared to 0.2% accounted

for by the poorest 10% of the population. This translates into an estimated Gini index of 0.69 (Statistics South Africa, 2008c). In this study the Gini index was estimated at 0.64. The difference (improvement) is a result of (i) adjustments made, in this study, in expenditure to account for underestimation of food and non-alcoholic beverages that represent a large part of poor households' expenditure and (ii) the use of an adult equivalence scale as opposed to per capita estimates in the Gini index. Given this high level of inequality, and the high prevalence of poverty noted earlier, it is important that health care payments do not impose undue hardship on the have-nots.

The results show that direct taxes are generally progressive health care financing mechanisms and as such impose a lower burden on the poor than on the rich. This, as noted earlier, is based on the structure of average and marginal tax rates. This progressivity is consistent in the literature, both in developed and developing countries. Similarly, financing health care through direct taxes induces a pro-poor income redistributive effect with minimal differential treatment. The progressivity effect accounts for most of the pro-poor income redistribution. However the case of indirect taxes is different. They were all found to be regressive and to induce pro-rich income redistribution. This result is not surprising because as noted earlier, indirect taxes are generally progressive health care financing mechanisms in developing countries (see Akazili *et al.*, 2011, Mtei *et al.*, 2012) but they are regressive in countries with more formalised economies (see Wagstaff *et al.*, 1999b, Smith, 2010).

There is a scope for improvement in indirect taxes (i.e. by making them less regressive than they are currently), especially by looking at the individual components such as value-added tax, fuel levy and excise taxes. For instance, in some African countries (for example Ghana and Tanzania) financing health care through value-added tax was progressive (Akazili *et al.*, 2011, Mtei *et al.*, 2012). Such progressivity in value-added tax was because items that are mainly consumed by poorer households are selectively exempted from VAT. In the case of South Africa, though there are items that are currently exempted from value-added tax, as noted earlier there is a need for an investigation to determine which goods, with the standard VAT rate, still take up a large fraction of poor households' expenditure and thereafter where possible exempt these from value-added tax. However this may lead to a loss in government tax revenue. With regard to excise taxes, especially alcohol taxes, the need to understand why the poor drink and bear a greater burden of this tax has been raised (Ataguba, 2012a) but needs

more research. This will enable the design of appropriate taxation mechanisms that will impose a lower burden on the poor than the current level.

The results of out-of-pocket payments as reported in Chapters 6 and 7 indicate that though they are a small proportion of total health care financing in South Africa, they were found to be mildly regressive with a pro-rich income redistributive effect. The increase in inequality associated with paying for health care out-of-pocket, as shown in Chapter 7, would have been 49% less if there were no differential treatment. This means that differential treatment alone accounted for over 49% of the rise in inequality over and above the base value predicted by the vertical component. About 27% of this differential treatment was due to the switching of ranks. The implication is that some 'old rich' were outranked by some 'new rich' in the process of redistribution and this can be deemed to be unfair. Similarly about 22% of the differential treatment is attributed to unequal treatment of equals. This means that individuals that have the same level of welfare before paying out-of-pocket for health care end up being treated unequally. While one may argue that out-of-pocket payments are as a result of deliberate choices individuals make and as such might not be considered as inequitable, in the context of South Africa this is unlikely to be the case because the bulk of such payments are co-payments made by those who are insured and as such are dictated by the decisions of their medical scheme. Many scheme members belong to a scheme selected by their employer and have limited choice (McIntyre, 2010b). The rest of the population that pay out-of-pocket for care do that as a result of the random incidence of illness. In fact for such people, these payments are generally described as to a large extent involuntary or non-discretionary (Wagstaff, 2009). Also as shown in the results of the decomposition in Chapter 7, the vertical component for out-of-pocket payments, though not statistically significant, is negative, signifying that the poor bear a greater burden. If the decision to make out-of-pocket payments were entirely voluntary, one would expect that it would be the rich who would be burdened with such payments. The issue here is that whether out-of-pocket payments in this context are defined as inequitable or simply unequal, it still remains possible to reduce the share of total health care financing that comes from out-of-pocket payments.

Turning to private health insurance, it was shown in Chapter 6 as a progressive financing source mainly because it is the rich who can afford to purchase cover and who therefore make contributions. As noted earlier, over 76% of the insured are located at the top quintile while less than 1% are in the bottom income quintile. However, within the top income decile, as

shown in Chapter 6, private health insurance was found to be regressive. This regressivity is because contributions to schemes are not income related (but are rather a flat amount), even though it is mainly the richer segment of the population that purchase insurance coverage. On top of this regressivity within the top income decile, privately insured households also face high co-payments at private sector facilities (McIntyre *et al.*, 2007). Also as shown in Chapter 7, differential treatment accounts for about 18% of the loss in redistributive power of the progressive private insurance contributions. This means that redistribution would have been higher (i.e. at its base value determined by the magnitude on the vertical effect) if there were no reranking or horizontal inequity. Reranking of individuals accounts for about 11% of the differential treatment due to households switching ranks after financing health care through private health insurance. Horizontal inequity was responsible for over 7% of the differential treatment. The question is: why are there still differential treatments in private health insurance? Is it due to the choice of different packages or the criteria for enrolment? It is the case that when insurance schemes enrol members by applying risk-rating there is a possibility for horizontal inequality and reranking (van Doorslaer *et al.*, 1999). However private health insurance schemes in South Africa, following the 1998 Medical Schemes Act are not permitted risk-rate but rather have to charge community rated contributions (McIntyre and Doherty, 2004). Therefore differential treatment is due to the choice of benefit package option and the number of household members enrolled. Because enrolment is not based on households but on individuals and premiums are a multiple of the number of household members enrolled, households who share the same level of welfare before contributions can enrol different numbers of individuals and face differential treatments. Also households with the same level of welfare before contributions can choose less comprehensive options to mitigate the impact of enrolling many household members. In this regard they may out-rank their 'equals'. However, the choice of low cost benefit options creates another form of inequity related to paying co-payments. All these points taken together indicate that the seemingly 'progressive' private health insurance still impose some burden on households, especially the 'poorer' insured members.

Overall, from the results presented in Chapters 6 and 7, the distribution of total health care financing was found to be fairer than the distribution of income. The poorest 40% of the population account for about 7.2% of total income and contribute about 2.8% of total health care finances compared to 83% of health care funding by the top 20% of the population. Total health care financing also reduces income inequality largely because it is progressive.

Looking at the three broad health care financing mechanisms – private health insurance, out-of-pocket payments and general taxes, what is the implication for progressivity and income redistribution associated with total health care financing of increased reliance on each? Would this make overall health care financing more or less progressive? Would it improve or reduce the income redistributive effect of total health care financing? Currently private health insurance alone accounts for over 45% of total health care financing. Expanding the coverage of private health insurance to a wider group, and maintaining its current features, could make private health insurance less progressive. It is likely to become a regressive health care financing mechanism if it is extended to a greater proportion of the population. This is because both the rich and the poor will belong to the schemes that charge relatively flat premiums. The regressive results within the top decile of the population as reported in Chapter 6 also point to this effect. Among the privately insured households, insurance contributions are regressive. This has been reported in countries such as the US and formerly Switzerland (Wagstaff *et al.*, 1999b) where insurance coverage is not restricted to the richer population groups. The implication of a regressive financing mechanism is a pro-rich income redistributive effect as reported in the US and formerly in Switzerland (van Doorslaer *et al.*, 1999). Therefore expanding private health insurance in its current form will not promote equity in health care financing.

The case of increased reliance on out-of-pocket payments in total health care financing is unpredictable. Currently such payments account for about 14% of total health care financing in South Africa. Based on international evidence, out-of-pocket payments have been reported as regressive or proportional (see for example Wagstaff *et al.*, 1999b, Cissé *et al.*, 2007, Abu-Zaineh *et al.*, 2008). They also induce a pro-rich income redistributive effect by increasing inequality in income (van Doorslaer *et al.*, 1999, Ichoku, 2006, Abu-Zaineh *et al.*, 2009). This inequity associated with out-of-pocket payments implies therefore that if all other things remain equal, increased reliance on this mechanism is unlikely to lead to an improvement in progressivity in overall health care financing. It will most likely lead to an increase in income inequality or a pro-rich income redistributive effect. Based on international evidence, increased reliance on out-of-pocket payments could impoverish the population or lead to financial catastrophe (Wagstaff and van Doorslaer, 2003, Knaul *et al.*, 2006, Ataguba, 2012b). Such increased reliance on out-of-pocket payments, in the case of South Africa, could arise from co-payments made by those who purchase private health insurance (McIntyre, 2010a).

They may also be due to patients, who do not purchase private health insurance, but utilise more and more private services if public health services are perceived as not meeting their expectations. In order to avoid such increases in out-of-pocket payments, it is important to pay particular attention to these issues in both the private and public health sectors.

General taxes on the other hand account for about 40% of total health care financing. International evidence has consistently shown that they are progressive and they induce pro-poor income redistributive effect. This is more so if direct taxes (that are consistently progressive) dominate indirect taxes. Based on international evidence on the progressivity and pro-poor income redistribution of general taxes, and evidence elsewhere (see for example Prakongsai *et al.*, 2007), the results presented in Chapters 6 and 7 imply that greater reliance on progressive general taxes will lead to a progressive overall health care financing system. This will also lead to a reduction in income inequality.

Based on the forgoing, increased reliance on general taxes is likely to keep total health care financing progressive. It is important to note here that such increased reliance does not necessarily imply increasing tax rates. For example, South Africa has witnessed “[e]xtensive tax reform and more efficient tax collection [*that*] have expanded revenue, permitting lower tax rates for both individuals and companies, and personal tax relief” (Ajam and Janine, 2007 p.745). Based on this and in the context of health care, increased reliance on general taxes could be achieved by sustaining the expansion in revenue collection, increasing the share of tax revenue allocated to the health sector, and sustaining GDP growth. If other financing mechanisms are used in addition to general taxes, it is important to avoid the negative impacts associated with them. Because indirect taxes are regressive, it is further important to ensure that these taxes do not impose additional burdens on the poorer groups relative to the richer groups.

The importance of general taxes in this context is important. They comprise a large share of public spending. The high level of income inequality in South Africa coupled with evidence that through increased public spending “...countries have succeeded in improving the access of the poor to basic social services while simultaneously realising significant increases in their incomes” (Squire, 1993) imply that there is a scope for improvement. Though there are debates as to the extent to which such spending will impact on the poor (Boadway and Marchand, 1998), it is the case that through fiscal discipline and ensuring an efficient system, it will likely have great impacts on the welfare of the poor. This follows from

evidence that the poor are more responsive to the price of social services (Squire, 1993). The essential role for increasing public spending through general taxes is for redistribution. Such “[r]edistribution to the poor in the form of improving their health, education, and nutrition is not only intrinsically important –in enhancing their capabilities to lead more fulfilling lives– but it is also instrumentally important in increasing their ‘human capital’ with lasting influence in the future” (Anand and Sen, 2000 p.2038).

The issue of prioritising social services including health care through increased public spending has been recognised in South Africa through the GEAR policy. It was noted that “[p]ublic sector reforms, comprising asset restructuring, budgetary reprioritisation and improved service delivery, underpin social and infrastructural development [...] and contribute to the redistribution of opportunities and income” (Department of Finance, 1994 p.7). In order to ensure such ‘redistribution of income and opportunities’ in South Africa, the progressive general taxes that induce pro-poor income redistribution hold promise for promoting an equitable health care financing system especially in the highly unequal country.

8.3 CONTRIBUTIONS OF THIS STUDY TO LITERATURE

This thesis makes both empirical and methodological contributions to the literature.

As previously noted, the assessment of equity in health care financing has been concentrated within developed countries (see van Doorslaer and Wagstaff, 1993). Recently however a few studies have emerged in developing countries including in Africa (see for example Ichoku, 2006, Cissé *et al.*, 2007, Akazili *et al.*, 2011). However such studies are either limited in their efforts to assess the extent of progressivity or only considered out-of-pocket payments (for example Ichoku, 2006). Within developing countries there is a strong argument for examining the overall distributional impact of health care financing, not least of all because of the extent of poverty but also the relative levels of socio-economic inequality and deprivation (Alkire and Santos, 2010). To date, however, nothing is known about the redistributive effect of overall health care financing in Africa. This study therefore fills this gap by investigating, for the first time, the redistributive effect of the entire health care financing system (taxes, out-of-pocket payments and private health insurance) in South Africa.

Three competing models for assessing redistributive effects (AJL, UL and DJA) were considered and compared in this thesis. Because South Africa presents a special case of a

highly unequal society (Gini index is in excess of 0.6), the application of the AJL and UL methodologies was particularly problematic largely due to the choice of close-equal groups. Even though a percentile approach was used to categorize close equal groups, it is argued that the main conclusion is that the DJA approach is more applicable under such circumstances of inequality because the results were consistent with *a priori* expectations.

Incidence assumptions for corporate income tax have been shown to be problematic in the literature. To date several assumptions have been used but there seems to be no consensus on what is most appropriate. In this thesis the Herfindahl Index, which shows market structure of industries in a country, was used to guide the incidence assumptions for corporate income tax in South Africa because the extent to which corporate tax is passed on to consumers depends on the existing market structure of industries. This index, if available for other countries, could prove helpful in refining the analysis of corporate tax incidence in the future.

This study also performed extensive sensitivity analyses to assess the distributional impact of health care financing. The major conclusion here is that the results are robust to the choice of parameters of aversion to inequality and reranking. The overall results are also robust to the choice of equivalence scale parameters used in generating adult equivalence.

There is a global call for countries to ensure equitable health care financing systems and also that their citizens have access to universal health care (World Health Organization, 2010). Achieving these goals requires *inter alia* considerable evidence which is unfortunately relatively scarce in Africa. Recently South Africa has embraced this call and is in the process of instituting a National Health Insurance with the goal of achieving universal coverage (Department of Health, 2011). This thesis provides empirical evidence on the extent of equity or inequity in the current South African health care financing system. Such evidence will assist in no small measure in shaping the current debate about health sector reforms in the country, especially those that relate to financing universal health system. This will also provide evidence for other African countries currently considering reforming their health care financing systems.

8.4 LIMITATIONS OF THE STUDY

This study has some limitations. One was the way corporate income tax was treated. Because as noted in Chapter 5, the literature on the treatment of corporate income tax has been divided, it was difficult to ascertain those who bear the final incidence of corporate income tax.

However, for the first time, the Herfindahl Index was used in order to allocate the burden of this tax. In addition, sensitivity analysis was performed by varying the relative shares of the burden between consumers and shareholders (or capital owners). It is also important to note that the dataset does not contain direct information on capital owners. In this case capital owners were identified as individuals who report dividends and contributions to investment.

Another limitation relates to the treatment of medical scheme contribution tax deductibility. The contributions to medical schemes should be deducted from gross personal income before applying the appropriate tax rates (and rebates) to extract total personal income tax payments. However it was not possible to do this in this thesis because medical scheme contributions were collected at the household level while personal income tax was computed at the individual level. It was noted that this may likely inflate, by a small proportion, the personal income tax liability computed for such individuals.

The analyses contained in this thesis assume away the barriers to assessing health care. This is especially the case with direct out-of-pocket payments. Also, it does not take into account the strategic decisions that individuals make in seeking health care. It only focuses on realised expenditure and does not include opportunity costs. This is however the internationally accepted standard for analysing the distributional impact of health care finance. Other forms of analyses may be used if the desire is to accommodate such issues.

The use of expenditure as a proxy for income in this study was necessitated by the lack of reliable data on income. In fact many households reported zero income rendering this variable unusable. The limitation encountered with the use of expenditure was the under-reporting of expenditures on some items such as food and non-alcoholic beverages. This was considered as one of the shortcomings of the IES 2005/06. However certain assumptions were made in order to adjust overall expenditure based on the proportion of under-reporting supplied by Statistics South Africa (the national statistical authority). Though the IES 2005/06 has some of these shortcomings, it is the only available comprehensive dataset with the range of variables needed for the kinds of analyses in this thesis. However, it is important to repeat these analyses on the IES 2010/11 that addresses the shortcomings of the IES 2005/06 which is scheduled to soon be released.

Assessing the distributional impact of health care financing using a snapshot of data (i.e. a given year) is informative but may not always reflect the pattern across years. This is another

limitation of the research. A dynamic approach to estimating the distributional impact of health care financing could be more illuminating. However the absence of reliable longitudinal data has limited the extent to which this can be achieved. It is hoped that when more recent data become available similar analyses can be performed. This limitation is particularly important if we are to assess the impact of reforms under the proposed National Health Insurance on the distribution of income in South Africa.

8.5 RECOMMENDATIONS FOR FUTURE RESEARCH

Equity in the health system can be better understood by looking at both the financing and delivery of health care. In this thesis only the distributional impact of health care financing has been analysed. This is insightful and important, but it is recommended that future research in this area can benefit from looking at both sides of the health system in order to give a more comprehensive picture of equity in the health system.

Another recommendation relates to the methodologies used in assessing equity in health care financing. The current methodologies used have been developed in the context of general taxes. They may not be particularly suitable for assessing fairness in all the forms of health care financing. The peculiarity of health care financing, especially private financing, requires modified methodologies that take into account the extent of population coverage and other factors to make conclusions about equity in financing. This is perhaps one of the reasons why it could be difficult (as this thesis has shown) to conclude that a progressive private health insurance is equitable in South Africa.

The analyses contained in this thesis rely on a partial equilibrium framework and thus ignore general equilibrium effects. Though the partial equilibrium framework is useful, it is recommended for future research that a more encompassing general equilibrium framework be used to explore, more fully, the distributional impact of health care finances in the general economy. This will answer the question: what is the impact of changing the pattern of health care financing on the overall economy?

This thesis decomposes the redistributive impact of specific health care financing sources in isolation of the others. Estimating the redistributive impact of the financing sources one by one runs the risk of not sufficiently taking into account the joint distribution of these sources across

the distribution of income. Future research should therefore consider estimating redistributive impact of health care financing to account for this joint distribution.

Because a health system is not static, it is further recommended that the distributional impact of health care financing should be monitored regularly with the availability of timely and reliable data. This is particularly the case when there are proposed changes in the health care financing system. This is the case with the current move towards ensuring universal health coverage in countries. It is therefore recommended that the distributional impact of all possible paths to ensuring universal coverage should be explored before actual implementation. This is to ensure that policies are based on informed evidence. In addition, there should also be monitoring and evaluation of the impact of reforms once implemented to assess whether the proposed objectives are actually being met.

The promotion of a progressive and equitable health care financing system is an imperative goal in many countries including South Africa. The results from this thesis point to general taxes as progressive and income redistributive. Based on these findings, the discussion contained in this thesis, and international evidence, the role of the government in the health sector is important. This role can be achieved through the reliance on progressive general taxes in financing health care. This will further promote an equitable health care financing system and contribute to an overall equitable universal health system.

REFERENCES

- Aaberge, R. & Melby, I. (1998). The sensitivity of income inequality to choice of equivalence scales. *Review of Income and Wealth*, 44, 565-569.
- Abu-Zaineh, M., Mataria, A., Luchini, S. & Jean-Paul, M. (2008). Equity in health care financing in Palestine: the value-added of the disaggregate approach. *Social Science & Medicine*, 66, 2308-2320.
- Abu-Zaineh, M., Mataria, A., Luchini, S. & Jean-Paul, M. (2009). Equity in health care finance in Palestine: the triple effects revealed. *Journal of Health Economics*, 28, 1071-1080.
- African Economic Outlook (2010a). *African Economic Outlook 2010*, Tunis, African Development Bank.
- African Economic Outlook (2010b). *African Economic Outlook 2011: South Africa*, Tunis, African Development Bank.
- Ahmad, E. & Stern, N. (1984). The theory of reform and Indian indirect taxes. *Journal of Public Economics*, 25, 259-298.
- Ajam, T. & Janine, A. (2007). Fiscal renaissance in a democratic South Africa. *Journal of African Economies*, 16, 745-781.
- Akazili, J. (2010). *Equity in health care financing in Ghana*, Cape Town, PhD thesis, University of Cape Town.
- Akazili, J., Gyapong, J. & McIntyre, D. (2011). Who pays for health care in Ghana? *International Journal for Equity in Health*, 10, 1-13.
- Alexander, L. & Schwarzschild, M. (1987). Liberalism, neutrality, and equality of welfare vs. equality of resources. *Philosophy and Public Affairs*, 16, 85-110.
- Alkire, S. & Santos, M. E. (2010). Acute multidimensional poverty: a new index for developing countries. *OPHI Working paper No. 38*, Oxford University.
- Amm, S. J. (2001). The incidence of corporate taxation - a suggested framework for measuring the incidence of South African corporate taxation. *An M.Comm dissertation submitted to the Department of Economics, University of Witwatersrand, Johannesburg*.
- Anand, S. (1977). Aspects of poverty in Malaysia. *Review of Income and Wealth*, 23, 1-16.
- Anand, S. & Hanson, K. (1997). Disability-adjusted life years: a critical review. *Journal of Health Economics*, 16, 685-702.
- Anand, S. & Sen, A. (2000). Human development and economic sustainability. *World Development*, 28, 2029-2049.

- Araar, A. & Duclos, J.-Y. (2009a). *DASP: Distributive Analysis Stata package*, Laval, PEP, CIRPÉE and World Bank, Université Laval.
- Araar, A. & Duclos, J.-Y. (2009b). *User manual for Stata package DASP: version 2.1*, Laval, PEP, World Bank, UNDP and Université Laval.
- Arndt, C. & Lewis, J. D. (2000). The macro implications of HIV/AIDS in South Africa: a preliminary assessment. *South African Journal of Economics*, 68, 856-887.
- Arneson, R. J. (1989). Equality and equal opportunity for welfare. *Philosophical Studies*, 56, 77-93.
- Arneson, R. J. (1990). Liberalism, distributive subjectivism, and equal opportunity for welfare. *Philosophy and Public Affairs*, 19, 158-194.
- Aronson, J. R., Johnson, P. & Lambert, P. J. (1994). Redistributive effect and unequal income tax treatment. *Economic Journal*, 104, 262-270.
- Aronson, J. R. & Lambert, P. J. (1994). Decomposing the Gini coefficient to reveal the vertical, horizontal, and reranking effects of income taxation. *National Tax Journal*, 47, 273-294.
- Ataguba, J. E. (2012a). Alcohol policy and taxation in South Africa: an examination of the economic burden of alcohol tax. *Applied Health Economics and Health Policy*, 10, 65-76.
- Ataguba, J. E. (2012b). Reassessing catastrophic health care payments with a Nigerian case study. *Health Economics, Policy and Law*, 7, 309-326.
- Ataguba, J. E., Akazili, J. & McIntyre, D. (2011). Socioeconomic-related health inequality in South Africa: evidence from General Household Surveys. *International Journal for Equity in Health*, 10, 48.
- Ataguba, J. E. & Goudge, J. (2012). The impact of health insurance on health-care utilisation and out-of-pocket payments in South Africa. *The Geneva Papers on Risk and Insurance-Issues and Practice*, 37, 633-654.
- Ataguba, J. E. & McIntyre, D. (2012). Who benefits from health services in South Africa? *Health Economics, Policy and Law*, DOI:10.1017/S1744133112000060.
- Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, 2, 244-63.
- Atkinson, A. B. (1973). How progressive should income tax be? In: Parkin, J. M. (ed.) *Essays in modern economics*. London: Longman.
- Atkinson, A. B. (1979). Horizontal equity and the distribution of the tax burden. In: Aaron, H. J. & Boskins, M. J. (eds.) *The economics of taxation*. Washington, DC.: The Brookings Institution.
- Atkinson, A. B. (1987). On the measurement of poverty. *Econometrica*, 55, 749-764.

- Auerbach, A. J. (2006). Who bears the corporate tax? A review of what we know. *In: Poterba, J. M. (ed.) Tax policy and the economy, volume 20.* Cambridge, MA: MIT Press.
- Auriol, E. & Warlters, M. (2005). Taxation base in developing countries. *Journal of Public Economics*, 89, 625-646.
- Banks, J. & Johnson, P. (1994). Equivalence scale relativities revisited. *The Economic Journal*, 104, 883-890.
- Barde, J. & Braathen, N. A. (2005). Environmentally related levies. *In: Cnossen, S. (ed.) Theory and practice of excise taxation: smoking, drinking, gambling, polluting, and driving.* New York: Oxford University Press.
- Bilger, M. (2008). Progressivity, horizontal inequality and reranking caused by health system financing: a decomposition analysis for Switzerland. *Journal of Health Economics*, 27, 1582-1593.
- Birch, S. & Donaldson, C. (2003). Valuing the benefits and costs of health care programmes: where's the 'extra' in extra-welfarism? *Social Science & Medicine*, 56, 1121-1133.
- Bird, R. M. & De Wulf, L. H. (1973). Taxation and income distribution in Latin America: a critical review of empirical studies. *International Monetary Fund Staff Papers*, 20, 639-682.
- Bishop, J. A., Chow, K. V. & Formby, J. P. (1994). Testing for marginal changes in income distributions with Lorenz and concentration curves. *International Economic Review*, 35, 479-488.
- Blackburn, R. M. & Jarman, J. (1993). Changing inequalities in access to British universities. *Oxford Review of Education*, 19, 197-215.
- Blackman, A., Osakwe, R. & Alpizar, F. (2010). Fuel tax incidence in developing countries: the case of Costa Rica. *Energy Policy*, 38, 2208-2215.
- Blackorby, C. & Donaldson, D. (1978). Measures of relative equality and their meaning in terms of social welfare. *Journal of Economic Theory*, 18, 59-80.
- Blundell, R. & Preston, I. (1995). Income, expenditure and the living standards of UK households. *Fiscal Studies*, 16, 40-54.
- Boadway, R. & Keen, M. (2000). Redistribution. *In: Atkinson, A. B. & Bourguignon, F. (eds.) Handbook of income distribution (Volume 1).* Amsterdam: North-Holland.
- Boadway, R. & Marchand, M. (1998). A review of public spending and the poor. *Review of Income and Wealth*, 44, 285-302.
- Borghi, J., Ataguba, J., Mtei, G., Akazili, J., Meheus, F., Rehnberg, C. & McIntyre, D. (2009). Methodological challenges in evaluating health care financing equity in data-poor contexts: lessons from Ghana, South Africa and Tanzania *In: Chernichovsky, D. & Hanson, K. (eds.) Innovations in health system finance in developing and transitional economies (Advances in health economics and health services research, volume 21).* Bingley, UK: Emerald Group Publishing.

- Bradshaw, D., Nannan, N., Laubscher, R., Groenewald, P., Joubert, J., Nojilana, B., Norman, R., Pieterse, D. & Schneider, M. (2006). *South African national burden of disease study 2000: estimates of provincial mortality: summary report*, Parow, Cape Town, South African Medical Research Council.
- Brouwer, W. B. F., Culyer, A. J., van Exel, N. & Rutten, F. F. H. (2008). Welfarism vs. extra-welfarism. *Journal of Health Economics*, 27, 325-338.
- Buehler, A. G. (1945). Ability to pay. *Tax Law Review*, 1, 243-258.
- Buhmann, B., Rainwater, L., Schmaus, G. & Smeeding, T. M. (1988). Equivalence scales, well-being, inequality, and poverty: sensitivity estimates across ten countries using the Luxembourg Income Study (LIS) database. *Review of Income and Wealth*, 34, 115-142.
- Burgess, R. & Stern, N. (1993). Taxation and development. *Journal of Economic Literature*, 31, 762-830.
- Burkhauser, R. V., Smeeding, T. M. & Merz, J. (1996). Relative inequality and poverty in Germany and the United States using alternative equivalence scales. *Review of Income and Wealth*, 42, 381-400.
- Cameron, A. C. & Trivedi, P. K. (2009). *Microeconometrics using Stata*, Texas, USA, Stata Press.
- Carrin, G., Evans, D. & Xu, K. (2007). Designing health financing policy towards universal coverage. *Bulletin of the World Health Organization*, 85, 652-652.
- Castano, R. A., Arbelaez, J. J., Giedion, U. B. & Morales, L. G. (2002). Equitable financing, out-of-pocket payments and the role of health care reform in Colombia. *Health Policy and Planning*, 17, 5-11.
- Cavagnero, E. & Bilger, M. (2010). Equity during an economic crisis: financing of the Argentine health system. *Journal of Health Economics*, 29, 479-488.
- Cissé, B., Luchini, S. & Moatti, J. P. (2007). Progressivity and horizontal equity in health care finance and delivery: what about Africa? *Health Policy*, 80, 51-68.
- Citro, C. F. & Michael, R. T. (1995). *Measuring poverty: a new approach*, Washington D.C, National Academy Press.
- Cleary, S., Silal, S., Birch, S., Carrara, H., Pillay-van Wyk, V., Rehle, T. & Schneider, H. (2011). Equity in the use of antiretroviral treatment in the public health care system in urban South Africa. *Health Policy*, 99, 261-266.
- Coast, J., Smith, R. D. & Lorgelly, P. (2008). Welfarism, extra-welfarism and capability: The spread of ideas in health economics. *Social Science & Medicine*, 67, 1190-1198.
- Cohen, G. A. (1989). On the currency of egalitarian justice. *Ethics*, 99, 906-944.

- Commission on Social Determinants of Health (2008). Closing the gap in a generation: health equity through action on the social determinants of health. Final Report of the Commission on Social Determinants of Health. Geneva: World Health Organization.
- Cook, P. J. & Moore, M. J. (1994). This tax's for you: The case for higher beer taxes. *National Tax Journal*, 47, 559-573.
- Coovadia, H., Jewkes, R., Barron, P., Sanders, D. & McIntyre, D. (2009). The health and health system of South Africa: historical roots of current public health challenges. *The Lancet*, 374, 817-834.
- Coulter, F. A. E., Cowell, F. A. & Jenkins, S. P. (1992). Equivalence scale relativities and the extent of inequality and poverty. *The Economic Journal*, 102, 1067-1082.
- Council for Medical Schemes (2006). Council for Medical Schemes annual report 2005-6. Pretoria: Council for Medical Schemes.
- Council for Medical Schemes (2010). Council for Medical Schemes annual report 2009-2010. Pretoria: Council for Medical Schemes.
- Crocker, L. (1977). Equality, solidarity, and Rawls' maximin. *Philosophy and Public Affairs*, 6, 262-66.
- Culyer, A. (1994). Commodities, characteristics of commodities, characteristics of people, utilities, and the quality of life. In: Baldwin, S., Godfery, C. & Propper, C. (eds.) *Quality of life: perspectives and policies*. London: Routledge.
- Culyer, A. J. (1989). The normative economics of health care finance and provision. *Oxford Review of Economic Policy*, 5, 34-58.
- Culyer, A. J. & Wagstaff, A. (1993). Equity and equality in health and health care. *Journal of Health Economics*, 12, 431-457.
- Cutler, D. M. & Katz, L. F. (1992). Rising inequality? Changes in the distribution of income and consumption in the 1980s. *American Economic Review*, 82, 546-551.
- Dardanoni, V. & Forcina, A. (1999). Inference for Lorenz curve orderings. *The Econometrics Journal*, 2, 49-75.
- Davidson, R. & Duclos, J.-Y. (1997). Statistical inference for the measurement of the incidence of taxes and transfers. *Econometrica*, 65, 1453-1465.
- de Wet, T. J. & van Heerden, J. H. (2003). The dividends from a revenue neutral tax on coal in South Africa. *South African Journal of Economic and Management Sciences*, 6, 473-497.
- Deaton, A. (1997). *The analysis of household surveys*, Baltimore, MD, The Johns Hopkins University Press.
- Deaton, A. (2002). Policy implications of the gradient of health and wealth. *Health Affairs*, 21, 13-30.

- Deaton, A. (2003). Health, inequality, and economic development. *Journal of Economic Literature*, 41, 113-158.
- Deaton, A. & Zaidi, S. (2002). Guidelines for constructing consumption aggregates for welfare analysis. *LSMS Working Paper No. 135*. Washington D.C.: World Bank Publications.
- Deaton, A. S. & Paxson, C. H. (1998). Aging and inequality in income and health. *The American Economic Review*, 88, 248-253.
- Denier, Y. (2007). *Efficiency, Justice and Care*, Dordrecht, Springer.
- Department of Finance (1994). *Growth, employment and redistribution: a macroeconomic strategy*, Pretoria, South African Department of Finance.
- Department of Health (1997). White Paper for the transformation of the health system in South Africa. Pretoria: South African Department of Health.
- Department of Health (2011). National Health Insurance in South Africa. Pretoria: National Department of Health, Republic of South Africa.
- Department of Health, Medical Research Council & OrcMacro (2007). *South Africa Demographic and Health Survey 2003*, Pretoria, Department of Health.
- Department of Labour (2010). Annual Labour Market Bulletin (April 2009 - March 2010). Pretoria: South African Department of Labour.
- Development Indicators (2010). Development indicators, 2010. Pretoria: The presidency, Republic of South Africa.
- Djolv, G. (2009). Methodological challenges in measuring business concentration as an official statistic: the case of the Hirschman-Herfindahl index. *A paper presented at the 2009 ISI conference*. Durban: International Statistical Institute.
- Dolan, P. (1999). Drawing a veil over the measurement of social welfare--a reply to Johannesson. *Journal of Health Economics*, 18, 387-390.
- Donaldson, C., Gerard, K., Mitton, C., Jan, S. & Wiseman, V. (2004). *Economics of health care financing: the visible hand*, Basingstoke, UK, Palgrave Macmillan.
- Donaldson, D. & Weymark, J. A. (1980). A single-parameter generalization of the Gini indices of inequality. *Journal of Economic Theory*, 22, 67-86.
- Donaldson, D. & Weymark, J. A. (1983). Ethically flexible Gini indices for Income distributions in the continuum. *Journal of Economic Theory*, 29, 353-358.
- du Toit, J. (2002). *The structure of the South African economy*, Johannesburg, The SA Financial Sector Forum.
- Duclos, J.-Y. (2000). Gini indices and the redistribution of income. *International Tax and Public Finance*, 7, 141-162.
- Duclos, J.-Y. & Araar, A. (2006). *Poverty and equity: measurement, policy, and estimation with DAD*, Berlin and Ottawa, Springer and IDRC.

- Duclos, J.-Y., Araar, A. & Fortin, C. (2010). DAD 4.6: A software for distributive analysis / analyse distributive. MIMAP programme, International Development Research Centre, Government of Canada and CREFA, Université Laval.
- Duclos, J.-Y., Jalbert, V. & Araar, A. (2003). Classical Horizontal Inequity and reranking: an Integrated approach. *In: Amiel, Y. & Bishop, J. A. (eds.) Fiscal policy, inequality and welfare (Research on economic inequality volume 10)*. Bingley: Emerald Group
- Duclos, J.-Y. & Lambert, P. J. (2000). A normative and statistical approach to measuring classical horizontal inequity. *Canadian Journal of Economics*, 33, 87-113.
- Dworkin, R. (1981a). What is equality? Part 1: equality of welfare. *Philosophy and Public Affairs*, 10, 185-246.
- Dworkin, R. (1981b). What is equality? Part 2: equality of resources. *Philosophy and Public Affairs*, 10, 283-345.
- Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association*, 171-185.
- Efron, B. & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, 1, 54-77.
- Elster, J. (1992). *Local justice: How institutions allocate scarce goods and necessary burdens*, New York, Russell Sage.
- Emran, M. S. & Stiglitz, J. E. (2005). On selective indirect tax reform in developing countries. *Journal of Public Economics*, 89, 599-623.
- Feldman, A. & Kirman, A. (1974). Fairness and envy. *The American Economic Review*, 64, 995-1005.
- Feldstein, M. (1973). On the optimal progressivity of the income tax. *Journal of Public Economics*, 2, 357-376.
- Feldstein, M. (1976). On the theory of tax reform. *Journal of Public Economics*, 6, 77-104.
- Filmer, D. & Pritchett, L. (2001). Estimating wealth effects without expenditure data—or tears: an application to educational enrollments in states of India. *Demography*, 38, 115-132.
- Fleurbaey, M. (2001). Egalitarian opportunities. *Law and Philosophy*, 20, 499-530.
- Foley, D. K. (1967). Resource allocation and the public sector. *Yale Economic Essays*, 7, 45-98.
- Formby, J., Seaks, T. & Smith, W. (1981). A comparison of two new measures of tax progressivity. *The Economic Journal*, 91, 1015-1019.
- Fourie, F. & Owen, A. (1993). Value added tax and regressivity in South Africa. *South African Journal of Economics*, 61, 308-319.

- Fuchs, V. R., Krueger, A. B. & Poterba, J. M. (1998). Economists' views about parameters, values, and policies: survey results in labor and public economics. *Journal of Economic Literature*, 36, 1387-1425.
- Fullerton, D. & Heutel, G. (2007). The general equilibrium incidence of environmental taxes. *Journal of Public Economics*, 91, 571-591.
- Fullerton, D. & Metcalf, G. E. (2002). Tax incidence. In: Auerbach, A. J. & Feldstein, M. (eds.) *Handbook of Public Economics*. North-Holland: Elsevier Science Publishers.
- Gandjour, A. & Lauterbach, K. W. (2003). Utilitarian theories reconsidered: common misconceptions, more recent developments, and health policy implications. *Health Care Analysis*, 11, 229-244.
- Garenne, M. (2004). Sex ratios at birth in populations of Eastern and Southern Africa. *Southern African Journal of Demography*, 9, 91-96.
- Gemmell, N. & Morrissey, O. (2005). Distribution and poverty impacts of tax structure reform in developing countries: how little we know. *Development Policy Review*, 23, 131-144.
- Gerard, K. & Mooney, G. (1993). QALY league tables: handle with care. *Health Economics*, 2, 59-64.
- Gerdtham, U.-G. & Sundberg, G. (1998). Redistributive effects of Swedish health care finance. *The International Journal of Health Planning and Management*, 13, 289-306.
- Gilson, L. & McIntyre, D. (2007). Post-apartheid challenges: household access and use of health care in South Africa. *International Journal of Health Services*, 37, 673-691.
- Glewwe, P. (1991). Household equivalence scales and the measurement of inequality: transfers from the poor to the rich could decrease inequality. *Journal of Public Economics*, 44, 211-216.
- Go, D. S., Kearney, M., Robinson, S. & Thierfelder, K. (2005). An analysis of South Africa's value added tax. *World Bank Policy Research Working Paper 3671*. Washington D.C.: World Bank.
- Gordon, R. & Li, W. (2009). Tax structures in developing countries: many puzzles and a possible explanation. *Journal of Public Economics*, 93, 855-866.
- Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, 80, 223-255.
- Gustafsson, B. & Uusitalo, H. (1990). The welfare state and poverty in Finland and Sweden from the mid 1960s to the mid 1980s. *Review of Income and Wealth*, 36, 249-266.
- Hansson, S. O. (2004). What are opportunities and why should they be equal? *Social Choice and Welfare*, 22, 305-316.
- Harberger, A. (1962). The incidence of the corporation income tax. *The Journal of Political Economy*, 70, 215-240.

- Harmon, C. & Nolan, B. (2001). Health insurance and health services utilization in Ireland. *Health Economics*, 10, 135-145.
- Hart, T. (1971). The inverse care law. *The Lancet*, 297, 405-412.
- Hines Jr, J. R. (2000). What is benefit taxation? *Journal of Public Economics*, 75, 483-492.
- Hirschman, A. O. (1964). The paternity of an index. *The American Economic Review*, 54, 761-762.
- Huang, Y. (1976). Distribution of the tax burden in Tanzania. *The Economic Journal*, 86, 73-86.
- Hurley, J. (1988). Welfarism, extra-welfarism and evaluative economic analysis in the health care sector. In: Barer, M. L., Getzen, T. E. & Stoddart, G. L. (eds.) *Health, health care and health economics: perspectives on distribution*. Chichester: John Wiley & Sons.
- Ichoku, H. E. (2006). *A distributional analysis of healthcare financing in a developing country: a Nigerian case study applying a decomposable Gini index*, Cape Town, PhD thesis, University of Cape Town.
- Ichoku, H. E. (2011). On the use of socioeconomic status indicators in the analysis of health inequalities and poverty in Africa. *Journal of International Development*, 23, 1004-1012.
- Ichoku, H. E., Fonta, W. M. & Ataguba, J. E. (2012). Political economy and history: making sense of health financing in sub-Saharan Africa. *Journal of International Development*, DOI: 10.1002/jid.2842.
- Jenkins, S. P. (1994). Social welfare function measures of horizontal inequity. In: Eichhorn, W. (ed.) *Model and methods of welfare and inequality*. Berlin: Springer Verlag.
- Jenkins, S. P. & Cowell, F. A. (1994). Parametric equivalence scales and scale relativities. *The Economic Journal*, 104, 891-900.
- Kahn, K. (2011). Population health in South Africa: dynamics over the past two decades. *Journal of Public Health Policy*, 32, S30-S36.
- Kahneman, D., Knetsch, J. L. & Thaler, R. H. (1986). Fairness and the assumptions of economics. *The Journal of Business*, 59, 285-300.
- Kakwani, N. (1980). *Income inequality and poverty: methods of estimation and policy applications*, Oxford, Oxford University Press.
- Kakwani, N., Wagstaff, A. & van Doorslaer, E. (1997). Socioeconomic inequalities in health: measurement, computation, and statistical inference. *Journal of Econometrics*, 77, 87-103.
- Kakwani, N. C. (1977). Measurement of tax progressivity: an international comparison. *Economic Journal*, 87, 71-80.

- Kakwani, N. C. (1984). On the measurement of tax progressivity and redistributive effect of taxes with applications to horizontal and vertical equity. *Advances in Econometrics*, 3, 149-168.
- Kale, R. (1995). Impressions of health in the new South Africa: a period of convalescence. *British Medical Journal*, 310, 1119-1122.
- Karoly, L. A. & Burtless, G. (1995). Demographic change, rising earnings inequality, and the distribution of personal well-being, 1959-1989. *Demography*, 32, 379-405.
- Khetan, C. P. & Poddar, S. N. (1976). Measurement of income tax progression in a growing economy: the Canadian experience. *Canadian Journal of Economics*, 9, 613-629.
- King, M. A. (1983). An index of inequality: with applications to horizontal equity and social mobility. *Econometrica*, 51, 99-115.
- Klasen, S. (1997). Poverty, inequality and deprivation in South Africa: an analysis of the 1993 SALDRU survey. *Social Indicators Research*, 41, 51-94.
- Klasen, S. & Woolard, I. (2009). Surviving unemployment without state support: unemployment and household formation in South Africa. *Journal of African Economies*, 18, 1-51.
- Klavus, J. (2001). Statistical inference of progressivity dominance: an application to health care financing distributions. *Journal of Health Economics*, 20, 363-377.
- Knaul, F., Arreola-Ornelas, H., Méndez-Carniado, O., Bryson-Cahn, C., Barofsky, J., Maguire, R., Miranda, M. & Sesma, S. (2006). Evidence is good for your health system: policy reform to remedy catastrophic and impoverishing health spending in Mexico. *The Lancet*, 368, 1828-1841.
- Konow, J. (2003). Which is the fairest one of all? A positive analysis of justice theories. *Journal of Economic Literature*, 41, 1188-1239.
- Kotlikoff, L. & Summers, L. (1987). Tax incidence. In: Auerbach, A. J. & Feldstein, M. (eds.) *Handbook of Public Economics*. North-Holland: Elsevier Science Publishers.
- Lairson, D. R., Hindson, P. & Hauquitz, A. (1995). Equity of health care in Australia. *Social Science & Medicine*, 41, 475-482.
- Lambert, P. J. (2001). *The distribution and redistribution of income*, Manchester, Manchester University Press.
- Lambert, P. J. & Ramos, X. (1997). Horizontal inequity and vertical redistribution. *International Tax and Public Finance*, 4, 25-37.
- Lamont, J. (2004). Distributive justice. In: F.Gaus, G. & Kukathas, C. (eds.) *Handbook of Political Theory*. London: Sage Publications Ltd.
- Lancaster, G., Ray, R. & Valenzuela, M. R. (1999). A cross-country study of equivalence scales and expenditure inequality on unit record household budget data. *Review of Income and Wealth*, 45, 455-82.

- Le Grand, J. (1987a). Equity, health, and health care. *Social Justice Research*, 1, 257-274.
- Le Grand, J. (1987b). Equity, well-being, and economic choice. *The Journal of Human Resources*, 22, 429-440.
- Le Grand, J. (1991). *Equity and choice: an essay in economics and applied philosophy*, London, Routledge.
- Lerman, R. I. & Yitzhaki, S. (1989). Improving the accuracy of estimates of Gini coefficients. *Journal of Econometrics*, 42, 43-47.
- Lerman, R. I. & Yitzhaki, S. (1995). Changing ranks and the inequality impacts of taxes and transfers. *National Tax Journal*, 48, 45-45.
- Limwattananon, S., Tangcharoensathien, V. & Prakongsai, P. (2007). Catastrophic and poverty impacts of health payments: results from national household surveys in Thailand. *Bulletin of the World Health Organization*, 85, 600-606.
- Logan, J. R. & Bian, Y. (1993). Inequalities in access to community resources in a Chinese city. *Social Forces*, 72, 555-576.
- Lorenz, M. O. (1905). Methods of measuring the concentration of wealth. *Publications of the American Statistical Association*, 9, 209-219.
- Lu, C., Chin, B., Lic, G. & Murray, C. J. L. (2009). Limitations of methods for measuring out-of-pocket and catastrophic private health expenditures. *Bulletin of the World Health Organization*, 87, 238-244.
- Lyon, A. B. & Schwab, R. (1995). Consumption taxes in a life-cycle framework: are sin taxes regressive? *The Review of Economics and Statistics*, 77, 389-340.
- Mabugu, R., Chitiga, M. & Amusa, H. (2009). The economic consequences of a fuel levy reform in South Africa. *South African Journal of Economic and Management Sciences*, 12, 280-296.
- Martinez-Vazquez, J. (2001). The impact of budgets on the poor: tax and benefit incidence. *International Studies Program, Georgia State University, Working Paper 01-10*. Georgia State University, Georgia.
- McClements, L. D. (1977). Equivalence scales for children. *Journal of Public Economics*, 8, 191-210.
- McGuire, A., Henderson, J. & Mooney, G. (1992). *The economics of health care: an introductory text*, New York, Routledge.
- McIntyre, D. (2010a). National health insurance: providing a vocabulary for public engagement. In: Padarath, A. & Fonn, S. (eds.) *South African Health Review*. Pretoria: Health Systems Trust.
- McIntyre, D. (2010b). National Health Insurance: Providing a vocabulary for public engagement. In: Fonn, S. & Padarath, A. (eds.) *South African Health Review 2010*. Durban: Health Systems Trust.

- McIntyre, D. (2010c). SHIELD Work Package 5 Report: modelling the estimated resource requirements of alternative health care financing reforms in South Africa. Health Economics Unit, University of Cape Town.
- McIntyre, D. (2012). What healthcare financing changes are needed to reach universal coverage in South Africa? *South African Medical Journal*, 102, 489-490.
- McIntyre, D. & Ataguba, J. E. (2012). Modelling the affordability and distributional implications of future health care financing options in South Africa. *Health Policy and Planning*, 27, i101-i112.
- McIntyre, D., Doherty, J. E. & Ataguba, J. E. (2012). Health care financing and expenditure: post-1994 progress and remaining challenges. In: Van Rensburg, H. C. J. (ed.) *Health and health care in South Africa*. Pretoria: Van Schaik.
- McIntyre, D., Garshong, B., Mtei, G., Meheus, F., Thiede, M., Akazili, J., Ally, M., Aikins, M., Mulligan, J. A. & Goudge, J. (2008). Beyond fragmentation and towards universal coverage: insights from Ghana, South Africa and the United Republic of Tanzania. *Bulletin of the World Health Organization*, 86, 871-876.
- McIntyre, D. & Gilson, L. (2002). Putting equity in health back onto the social policy agenda: experience from South Africa. *Social Science & Medicine*, 54, 1637-1656.
- McIntyre, D., Goudge, J., Harris, B., Nxumalo, N. & Nkosi, M. (2009). Prerequisites for National Health Insurance in South Africa: results of a national household survey. *South African Medical Journal*, 99, 725-729.
- McIntyre, D., Muirhead, D. & Gilson, L. (2002). Geographic patterns of deprivation in South Africa: informing health equity analyses and public resource allocation strategies. *Health Policy and Planning*, 17, 30-39.
- McIntyre, D., Thiede, M., Dahlgren, G. & Whitehead, M. (2006). What are the economic consequences for households of illness and of paying for health care in low-and middle-income country contexts? *Social Science & Medicine*, 62, 858-865.
- McIntyre, D., Thiede, M., Nkosi, M., Mutyambizi, V., Castillo-Riquelme, M., Gilson, L., Erasmus, E. & Goudge, J. (2007). A critical analysis of the current South African health system. Cape Town: Health Economics Unit, University of Cape Town and Centre for Health Policy, University of the Witwatersrand.
- McIntyre, D. E. & Doherty, J. E. (2004). Health care financing and expenditure: progress since 1994 and remaining challenges. In: van Rensburg, H. C. J. (ed.) *Health and health care in South Africa*. Pretoria: Van Schaik.
- McLeod, H. & Ramjee, S. (2007). Medical schemes. In: Harrison, S., Bhana, R. & Ntuli, A. (eds.) *South African Health Review 2007*. Durban: Health Systems Trust.
- McLure, C. E. (1975). General equilibrium incidence analysis: the Harberger model after ten years. *Journal of Public Economics*, 4, 125-161.

- MDG Report (2011). *Assessing progress in Africa toward the Millennium Development Goals*, Denmark, African Development bank, Economic Commission for Africa, African Union, United Nations Development Programme.
- Miller, G. J. (2005). Government fiscal policy impacts. *In: Robbins, D. (ed.) Handbook of Public Sector Economics*. Boca Raton, FL: Marcel Dekker/CRC Press LLC.
- Mooney, G. (1989). QALYs: are they enough? A health economist's perspective. *Journal of medical ethics*, 15, 148-152.
- Mooney, G. (1996). And now for vertical equity? Some concerns arising from aboriginal health in Australia. *Health Economics*, 5, 99-103.
- Mooney, G. (2000). Vertical equity in health care resource allocation. *Health Care Analysis*, 8, 203-215.
- Mooney, G. (2003). Inequity in Australian health care: how do we progress from here? *Australian and New Zealand Journal of Public Health*, 27, 267-270.
- Mooney, G. (2009). *Challenging health economics*, New York, Oxford University Press.
- Mooney, G. & Gilson, L. (2009). The economic situation in South Africa and health inequities: a comment. *The Lancet*, 374, 858-859.
- Mooney, G. & Jan, S. (1997). Vertical equity: weighting outcomes? or establishing procedures? *Health Policy*, 39, 79-87.
- Mooney, G. H. (1994). *Key issues in health economics*, New York, Harvester Wheatsheaf.
- Moulin, H. (2003). *Fair division and collective welfare*, London, England, The MIT Press.
- Moulin, H. & Roemer, J. (1989). Public ownership of the external world and private ownership of self. *The Journal of Political Economy*, 97, 347-367.
- Mtei, G. & Borghi, J. (2010). SHIELD Work Package 2 Report: an assessment of health care financing progressivity in Tanzania. Health Economics Unit, University of Cape Town.
- Mtei, G., Makawia, S., Ally, M., Kuwawenaruwa, A., Meheus, F. & Borghi, J. (2012). Who pays and who benefits from health care? An assessment of equity in health care financing and benefit distribution in Tanzania. *Health Policy and Planning*, 27, i23-i34.
- Murray, C. J. (1994). Quantifying the burden of disease: the technical basis for disability-adjusted life years. *Bulletin of the World Health Organization*, 72, 429-445.
- Musgrave, R. A. (1985). A brief history of fiscal doctrine. *In: Auerbach, A. J. & Feldstein, M. (eds.) Handbook of Public Economics*. North-Holland: Elsevier Science Publishers.
- Musgrave, R. A. (1996). Progressive taxation, equity, and tax design. *In: Slemrod, J. (ed.) Tax progressivity and income inequality*. New York: Cambridge University Press.
- Musgrave, R. A., Carroll, J. J., Cook, L. D. & Frane, L. (1951). Distribution of tax payments by income groups: a case study for 1948. *National Tax Journal*, 4, 1-53.

- Musgrave, R. A. & Thin, T. (1948). Income tax progression, 1929-48. *The Journal of Political Economy*, 56, 498-514.
- Mwenge, F. (2010). *Progressivity and determinants of out of pocket health care financing in Zambia*, MPH thesis, University of Cape Town.
- Narayan, D., Patel, R., Schafft, K., Rademacher, A. & Koch-Schulte, S. (2000). *Voices of the poor: can anyone hear us?*, New York, Oxford University Press.
- Narveson, J. (1983a). On Dworkinian equality. *Social Philosophy and Policy*, 1, 1-23.
- Narveson, J. (1983b). Reply to Dworkin. *Social Philosophy and Policy*, 1, 41-44.
- National Treasury (2004). 2004 Budget Review. Pretoria, National Treasury.
- National Treasury (2005). 2005 Budget Review. Pretoria, National Treasury.
- National Treasury (2006). 2006 Budget Review. Pretoria, National Treasury.
- National Treasury (2007). 2007 Budget Review. Pretoria, National Treasury.
- National Treasury (2008). 2008 Budget Review. Pretoria, National Treasury.
- National Treasury (2009a). 2009 Budget Review. Pretoria, National Treasury.
- National Treasury (2009b). *Provincial Budgets and Expenditure Review*, Pretoria, National Treasury.
- National Treasury (2010). 2010 Budget Review. Pretoria, National Treasury.
- National Treasury (2011). 2011 Budget Review. Pretoria, National Treasury.
- Nelson, J. A. (1993). Household equivalence scales: theory versus policy? *Journal of Labor Economics*, 11, 471-493.
- Nelson Jr, W. R. (2001). Incorporating fairness into game theory and economics: comment. *American Economic Review*, 91, 1180-1183.
- Nozick, R. (1973). Distributive justice. *Philosophy and Public Affairs*, 3, 45-126.
- Nozick, R. (1974). *Anarchy, state, and utopia*, New York, Basic Books.
- O'Donnell, O., van Doorslaer, E., Rannan-Eliya, R. P., Somanathan, A., Adhikari, S. R., Akkazieva, B., Harbianto, D., Garg, C. C., Hanvoravongchai, P., Herrin, A. N., Huq, M., Ibragimova, S., Karan, A., Kwon, S., Leung, G., Lu, J., Ohkusa, Y., Pande, B., Racelis, R., Tin, K., Tisayaticom, K., Trisnantoro, L., Wan, Q., Yang, B. & Zhao, Y. (2008a). Who pays for health care in Asia? *Journal of Health Economics*, 27, 460-475.
- O'Donnell, O., Van Doorslaer, E., Wagstaff, A. & Lindelow, M. (2008b). *Analyzing health equity using household survey data: a guide to techniques and their implementation*, Washington DC., World Bank Publications.

- O'Donoghue, T. & Rabin, M. (2006). Optimal sin taxes. *Journal of Public Economics*, 90, 1825-1849.
- Olsen, J. A. (1997). Theories of justice and their implications for priority setting in health care. *Journal of Health Economics*, 16, 625-639.
- Oosthuizen, M. (2008). Estimating poverty lines for South Africa. *Discussion document: commissioned by the Department of Social Development*. Pretoria: South African Government Information.
- Pazner, E. A. & Schmeidler, D. (1974). A difficulty in the concept of fairness. *The Review of Economic Studies*, 41, 441-443.
- Pechman, J. A. & Okner, B. A. (1974). *Who bears the tax burden?*, Washington D.C., Brookings Institution.
- Pirttilä, J. & Uusitalo, R. (2010). A 'leaky bucket' in the real world: estimating inequality aversion using survey data. *Economica*, 77, 60-76.
- Plotnick, R. (1981). A measure of horizontal inequity. *The Review of Economics and Statistics*, 63, 283-288.
- Porath, E. B. & Gilboa, I. (1994). Linear measures, the Gini index, and the income-equality trade-off. *Journal of Economic Theory*, 64, 443-467.
- Poterba, J. M. (1991). Is the gasoline tax regressive? In: Bradford, D. (ed.) *Tax policy and the economy*. Boston: MIT Press.
- Prakongsai, P., Tangcharoensathien, V. & Tisayatikom, K. (2007). Who benefit from government health spending before and after universal coverage in Thailand? *Journal of Health Sciences*, 16, S20-S36.
- Prozzi, J. P., Naude, C. & Sperling, D. (2002). Transportation in developing countries: greenhouse gas scenarios for South Africa. Arlington: Pew center on global climate change.
- Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American Economic Review*, 83, 1281-1302.
- Rawls, J. (1971). *A theory of justice*, Harvard, MA, Harvard University Press.
- Rawls, J. (1974). Some reasons for the maximin criterion. *American Economic Review*, 64, 141-146.
- Rawls, J. (1999). *A theory of justice: revised edition*, Cambridge, MA., Harvard University Press.
- Republic of South Africa (2004). National Health Act No. 61 of 2003. Pretoria: The Presidency.
- Republic of South Africa (2010). *Millennium Development Goals country report 2010*, Pretoria, South African Government.

- Reynolds, M. & Smolensky, E. (1977). *Public expenditures, taxes, and the distribution of income: the United States 1950, 1961, 1970*, New York, Academic Press.
- Ringen, S. (1991). Households, standard of living, and inequality. *Review of Income and Wealth*, 37, 1-13.
- Robilliard, A. S. & Robinson, S. (2003). Reconciling household surveys and national accounts data using a cross entropy estimation method. *Review of Income and Wealth*, 49, 395-406.
- Roemer, J. E. (1996). *Egalitarian perspectives: essays in philosophical economics*, New York, Cambridge University Press.
- Rosenbluth, G. (1955). Measures of concentration. *Business concentration and price policy*. Princeton: National Bureau Committee for Economic Research.
- Runciman, W. G. (1966). *Relative deprivation and social justice: a study of attitudes to social inequality in twentieth-century England, Berkeley and Los Angeles*, Berkeley CA, University of California Press.
- Runciman, W. G. & Garrison, W. (1972). *Relative deprivation and social justice*, Routledge & Kegan Paul.
- Sahn, D. E. & Stifel, D. (2003). Exploring alternative measures of welfare in the absence of expenditure data. *Review of Income and Wealth*, 49, 463-489.
- Sahn, D. E., Younger, S. D. & Simler, K. R. (2000). Dominance testing of transfers in Romania. *Review of Income and Wealth*, 46, 309-327.
- Scarre, G. (1996). *Utilitarianism*, New York, Routledge.
- Sen, A. (1973). *On economic inequality*, Oxford, Oxford University Press.
- Sen, A. (1976). Poverty: an ordinal approach to measurement. *Econometrica*, 44, 219-231.
- Sen, A. (1977). Social choice theory: a re-examination. *Econometrica*, 45, 53-89.
- Sen, A. (1979). Utilitarianism and welfarism. *Journal of Philosophy*, 76, 463-489.
- Sen, A. (1984). The living standard. *Oxford Economic Papers*, 36, 74-90.
- Sen, A. (1985). *Commodities and capabilities*, Amsterdam, North Holland.
- Sen, A. (1988). *On ethics and economics*, Oxford, UK, Blackwell.
- Sen, A. (2009). *The idea of justice*, Cambridge, Massachusetts, Belknap Press.
- Shah, A. & Whalley, J. (1991). Tax incidence analysis of developing countries: an alternative view. *The World Bank Economic Review*, 5, 535-552.
- Shorrocks, A. (1978). Income inequality and income mobility. *Journal of Economic Theory*, 19, 376-393.

- Simkins, C., Woolard, I. & Thompson, K. (2000). An analysis of the burden of taxes in the South African economy for the years between 1995/96 and 1999/2000. Pretoria: South African Department of Finance.
- Slemrod, J. (1996). Introduction. *In: Slemrod, J. (ed.) Tax progressivity and income inequality*. New York: Cambridge University Press.
- Slemrod, J. & Bakija, J. M. (2004). *Taxing ourselves: a citizen's guide to the debate over taxes*, Cambridge, MA, MIT Press.
- Smith, S. (2010). Equity in Irish health care financing: measurement issues. *Health Economics, Policy and Law*, 5, 149-169.
- Squire, L. (1993). Fighting poverty. *The American Economic Review*, 377-382.
- StataCorp (2009). *Stata: release 11 - Statistical software*, College Station, TX, StataCorp LP.
- Statistics South Africa (2008a). Explanatory note on the 2006 consumer price index weights Pretoria: Statistics South Africa.
- Statistics South Africa (2008b). Income and expenditure of households 2005/2006 metadata (version 2). Pretoria: Statistics South Africa.
- Statistics South Africa (2008c). Income and expenditure of households 2005/2006: analysis of results. Pretoria: Statistics South Africa.
- Statistics South Africa (2008d). Income and expenditure of households 2005/2006: Statistical release P0100. Pretoria: Statistics South Africa.
- Statistics South Africa (2011). *Mid-year population estimates 2011*, Pretoria, Statistics South Africa.
- Stiglitz, J. E. (1982). Utilitarianism and horizontal equity: the case for random taxation. *Journal of Public Economics*, 18, 1-33.
- Suits, D. B. (1977). Measurement of tax progressivity. *American Economic Review*, 67, 747-752.
- Szende, A. & Culyer, A. J. (2006). The inequity of informal payments for health care: The case of Hungary. *Health Policy*, 75, 262-271.
- Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, 1, 39-60.
- Ugá, M. A. D. & Santos, I. S. (2007). An analysis of equity in Brazilian health system financing. *Health Affairs*, 26, 1017.
- UNDP (2011). *Human Development Report - sustainability and equity: a better future for all*, New York, Palgrave Macmillan.
- Urban, I. & Lambert, P. J. (2008). Redistribution, horizontal inequity, and reranking: how to measure them properly. *Public Finance Review*, 36, 563-587.

- van de Ven, J., Creedy, J. & Lambert, P. J. (2001). Close equals and calculation of the vertical, horizontal and reranking effects of taxation. *Oxford Bulletin of Economics and Statistics*, 63, 381-394.
- van Doorslaer, E. & Wagstaff, A. (1993). Equity in the finance of health care: methods and findings. In: van Doorslaer, E., Wagstaff, A. & Rutten, F. (eds.) *Equity in the finance and delivery of health care: an international perspective*. Oxford: Oxford University Press.
- van Doorslaer, E., Wagstaff, A., van der Burg, H., Christiansen, T., Citoni, G., Di Biase, R., Gerdtham, U. G., Gerfin, M., Gross, L. & Häkinnen, U. (1999). The redistributive effect of health care finance in twelve OECD countries. *Journal of Health Economics*, 18, 291-313.
- van Zon, A. & Muysken, J. (2005). Health as a principal determinant of economic growth. In: López-Casasnovas, G., Rivera, B. & Currais, L. (eds.) *Health and economic growth: findings and policy implications*. Cambridge, MA: The MIT press.
- Varian, H. (1974). Equity, envy, and efficiency. *Journal of Economic Theory*, 9, 63-91.
- Varian, H. R. (1975). Distributive justice, welfare economics, and the theory of fairness. *Philosophy and Public Affairs*, 4, 223-247.
- Wagle, U. (2008). *Multidimensional poverty measurement: concepts and applications*, New York, USA, Springer Publishers.
- Wagstaff, A. (1991). QALYs and the equity-efficiency trade-off. *Journal of Health Economics*, 10, 21-41.
- Wagstaff, A. (2002a). Poverty and health sector inequalities. *Bulletin of the World Health Organization*, 80, 97-105.
- Wagstaff, A. (2002b). Reflections on and alternatives to WHO's fairness of financial contribution index. *Health Economics*, 11, 103-115.
- Wagstaff, A. (2008). Measuring financial protection in health. *World Bank Policy Research Working Paper 4554*. Washington, D.C.: World Bank.
- Wagstaff, A. (2009). Measuring financial protection in health. In: Smith, P. C., Mossialos, E., Papanicolas, I. & Leatherman, S. (eds.) *Performance measurement for health system improvement: experiences, challenges and prospects*. New York: Cambridge University Press.
- Wagstaff, A. & van Doorslaer, E. (1992). Equity in the finance of health care: some international comparisons. *Journal of Health Economics*, 11, 361-387.
- Wagstaff, A. & van Doorslaer, E. (1993). Equity in the finance of health care: concepts and definitions. In: van Doorslaer, E., Wagstaff, A. & Rutten, F. (eds.) *Equity in the finance and delivery of health care: an international perspective*. Oxford: Oxford University Press.

- Wagstaff, A. & van Doorslaer, E. (1997). Progressivity, horizontal equity and reranking in health care finance: a decomposition analysis for the Netherlands. *Journal of Health Economics*, 16, 499-516.
- Wagstaff, A. & van Doorslaer, E. (2000). Equity in health care finance and delivery. In: Culyer, A. J. & Newhouse, J. P. (eds.) *Handbook of Health Economics*. Amsterdam: North-Holland.
- Wagstaff, A. & van Doorslaer, E. (2001). Paying for health care: quantifying fairness, catastrophe and impoverishment, with applications to Vietnam, 1993-98. *World Bank Policy Research Working Paper 2715*. Washington D.C.: World Bank.
- Wagstaff, A. & van Doorslaer, E. (2003). Catastrophe and impoverishment in paying for health care: with applications to Vietnam 1993-98. *Health Economics*, 12, 921-33.
- Wagstaff, A., van Doorslaer, E. & Paci, P. (1989). Equity in the finance and delivery of health care: some tentative cross-country comparisons. *Oxford Review of Economic Policy*, 5, 89-112.
- Wagstaff, A., van Doorslaer, E., van der Burg, H., Calonge, S., Christiansen, T., Citoni, G., Gerdtham, U. G., Gerfin, M., Gross, L. & Häkkinen, U. (1999a). Redistributive effect, progressivity and differential tax treatment: personal income taxes in twelve OECD countries. *Journal of Public Economics*, 72, 73-98.
- Wagstaff, A., van Doorslaer, E., van der Burg, H., Calonge, S., Christiansen, T., Citoni, G., Gerdtham, U. G., Gerfin, M., Gross, L., Häkkinen, U., Johnson, P., John, J., Klavus, J., Lachaud, C., Lauritsen, J., Leu, R., Nolan, B., Perán, E., Pereira, J., Propper, C., Puffer, F., Rochaix, L., Rodríguez, M., Schellhorn, M., Sundberg, G. & Winkelhake, O. (1999b). Equity in the finance of health care: some further international comparisons. *Journal of Health Economics*, 18, 263-290.
- Walls, M. & Hanson, J. (1999). Distributional aspects of an environmental tax shift: the case of motor vehicle emissions taxes. *National Tax Journal*, 52, 53-65.
- Weeks, J. (1999). Stuck in low GEAR? Macroeconomic policy in South Africa, 1996-98. *Cambridge Journal of Economics*, 23, 795-811.
- Weinstein, M. C. & Manning Jr, W. (1997). Theoretical issues in cost-effectiveness analysis. *Journal of Health Economics*, 16, 121-128.
- Woolard, I. D. & Leibbrandt, M. (1999). Measuring poverty in South Africa. DPRU working paper No 99/33: DPRU, University of Cape Town.
- World Bank (2011a). *Africa Development Indicators (FACTOIDS)*, Washington D.C., World Bank.
- World Bank (2011b). World Development Indicators database. Washington D.C.: World Bank.
- World Health Organization (2000). *The World Health Report 2000 - health systems: improving performance*, Geneva, World Health Organization.

- World Health Organization (2001). *Macroeconomics and health: investing in health for economic development*, Geneva, World Health Organisation.
- World Health Organization (2010). *The World Health Report 2010: Health systems financing: the path to universal coverage*, Geneva, World Health Organization.
- World Health Organization (2011). *World Health Statistics 2011*, Geneva, World Health Organisation.
- Xu, K., Evans, D. B., Kawabata, K., Zeramdini, R., Klavus, J. & Murray, C. J. L. (2003). Household catastrophic health expenditure: a multicountry analysis. *The Lancet*, 362, 111-117.
- Yaari, M. E. (1988). A controversial proposal concerning inequality measurement. *Journal of Economic Theory*, 44, 381-397.
- Yitzhaki, S. (1983). On an extension of the Gini inequality index. *International Economic Review*, 24, 617-628.
- Younger, S. D., Sahn, D. E., Haggblade, S. & Dorosh, P. A. (1999). Tax incidence in Madagascar: an analysis using household data. *The World Bank Economic Review*, 13, 303-331.
- Yu, C. P., Whynes, D. K. & Sach, T. H. (2008). Equity in health care financing: the case of Malaysia. *International Journal for Equity in Health*, 7, e15.
- Yu, C. P., Whynes, D. K. & Sach, T. H. (2011). Reform towards national health insurance in Malaysia: the equity implications. *Health Policy*, 100, 256-263.